

# Interpretatie van diagnostische gegevens : theoretische en praktische bijdragen aan de diagnostische epidemiologie in de extramurale gezondheidszorg

Citation for published version (APA):

Knottnerus, J. A. (1986). *Interpretatie van diagnostische gegevens : theoretische en praktische bijdragen aan de diagnostische epidemiologie in de extramurale gezondheidszorg*. [Maastricht University]. Rijksuniversiteit Limburg. <https://doi.org/10.26481/dis.19861212jk>

## Document status and date:

Published: 01/01/1986

## DOI:

[10.26481/dis.19861212jk](https://doi.org/10.26481/dis.19861212jk)

## Document Version:

Publisher's PDF, also known as Version of record

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

## Take down policy

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

Download date: 05 May. 2023

# **Interpretatie van diagnostische gegevens**

Theoretische en praktische bijdragen aan de diagnostische epidemiologie in de extramurale gezondheidszorg

## **Proefschrift**

Ter verkrijging van de graad van doctor in de geneeskunde aan de Rijksuniversiteit Limburg te Maastricht, op gezag van de Rector Magnificus, Prof. Dr. F.I.M. Bonke, volgens het besluit van het College van Dekanen in het openbaar te verdedigen op vrijdag 12 december 1986 om 14.00 uur.

door

André Knottnerus  
geboren te  
Nieuw-Beerta

# အသံသယနှင့် အကျိုးသက်ရောက်မှုကို တွေးခေါ်ခြင်းများ

အသံသယသည် အကျိုးသက်ရောက်မှုကို တွေးခေါ်ခြင်းများကို ဖြစ်ပေါ်စေသည်။

အသံသယ

အသံသယသည် အကျိုးသက်ရောက်မှုကို တွေးခေါ်ခြင်းများကို ဖြစ်ပေါ်စေသည်။ အသံသယသည် အကျိုးသက်ရောက်မှုကို တွေးခေါ်ခြင်းများကို ဖြစ်ပေါ်စေသည်။ အသံသယသည် အကျိုးသက်ရောက်မှုကို တွေးခေါ်ခြင်းများကို ဖြစ်ပေါ်စေသည်။

အသံသယ

အသံသယသည် အကျိုးသက်ရောက်မှုကို တွေးခေါ်ခြင်းများကို ဖြစ်ပေါ်စေသည်။

Promotoren: Prof. Dr. P.G. Knipschild  
Prof. Dr. F. Sturmans

Referenten: Prof. Dr. C.A. de Geus  
Prof. Dr. H.G.M. van der Velden

© 1986: J.A. Knottnerus, Maastricht

Omslagtekening: Bart Knottnerus





INHOUDSOPGAVE.

	pagina
1. Inleiding	4

Deel I: Interpretatie van diagnostische gegevens en de verschillen tussen eerste- en tweedelijns geneeskunde.

2. Interpretatie van diagnostische gegevens, een onontgonnen terrein in de huisartsgeneeskunde. gepubliceerd in: Huisarts en Wetenschap 1983; 26: 363-368. en in: Journal of the Royal College of General Practitioners 1985; 35: 270-274.	12
3. Symptomatologie en selectiebias. Vertekening van het verband tussen klachten en diagnosen ten gevolge van selectie naar hogere echelons. gepubliceerd in: Huisarts en Wetenschap 1985; 28: 325-330.	19

Deel II: Enkele observationele onderzoeken naar het verband tussen klachten en diagnostische bevindingen.

4. Hoofdpijn en hoge bloeddruk. gepubliceerd in: Huisarts en Wetenschap 1984; 27: 441-445.	26
5. Hemoglobine en 'anemische' klachten zoals moeheid. gepubliceerd in: Tijdschrift voor Sociale Gezondheidszorg 1983; 61: 788-792.	31
6. Onverklaarde moeheid en hemoglobinegehalte; een onderzoek vanuit de huisartspraktijk. gepubliceerd in: Nederlands Tijdschrift voor Geneeskunde 1986; 130: 402-405. en, in uitgebreidere vorm, in: The Canadian Family Physician 1986; 32: 1601-1604.	36
7. Serumferritine bij moeheidspatiënten en gezonde controlepersonen. aanvaard voor publicatie in het Nederlands Tijdschrift voor Geneeskunde.	40
8. Het hemoglobinegehalte van zwangeren en subjectieve klachten.	49
9. Diagnostische conclusies van de huisarts naar aanleiding van onverklaarde moeheid. aanvaard voor publicatie in Huisarts en Wetenschap.	67
10. Klachten op de borst: omgaan met onzekerheden. gepubliceerd in: Huisarts en Wetenschap 1985; 28: 159-164.	78

Deel III: Medische besliskunde, concepten en toepassingen.

11. Principes van besliskunde. 86  
gepubliceerd in: Nieuw Kompas voor de huisarts juli 1984  
II/ 4: 1-15.
12. De zeeffunctie van de huisarts, het totale plaatje is van 101  
belang.  
gepubliceerd in: The Practitioner (Nederlandse uitgave)  
1985; 2: 1251-1254.
13. Diagnostiek van de palpabele lymfeklier door huisarts en 105  
specialist, een besliskundige benadering.  
gepubliceerd in: the Practitioner (Nederlandse uitgave)  
1985; 2: 1085-1097.
14. Diagnostische hypothesen en a priori kansen in de huis- 114  
artsgeneeskunde.  
gepubliceerd in: Huisarts en Wetenschap 1986; 29: 269-273.
15. Moeheid: verschillen in benadering tussen huisartsen en 119  
internisten.  
gepubliceerd in: Medisch Contact 1986; 41: 54-57.
16. Intercollegiale vergelijking. Oefenen met huisartsen in 123  
opleiding.  
gepubliceerd in: Medisch Contact 1986; 41: 574-577.
17. Het belang van therapeutische mogelijkheden voor diagnos- 127  
tische beslissingen.  
gepubliceerd in: The Practitioner (Nederlandse uitgave)  
1984; 1: 769-777.
18. De asymptomatische carotisstenose, wel of niet opereren? 133  
Een besliskundige benadering.  
aanvaard voor publicatie in het Nederlands Tijdschrift  
voor Geneeskunde.
19. Over vertekening van testuitslagen door reeds aanwezige 151  
informatie.  
gepubliceerd in: The Practitioner (Nederlandse uitgave)  
1985; 2: 287-295.

Deel IV: Diagnostische tests nader bezien.

20. Diagnostische tests. Normen, referentiewaarden en test- 160  
kenmerken nader bezien: een inventarisatie.
21. De invloed van verwijsgedrag op het onderscheidend ver- 209  
mogen van diagnostische tests.

**Deel V : Enkele observationele onderzoeken naar het verband tussen uitslagen van diagnostische tests en prognose.**

22. Hemoglobine en zwangerschapsuitkomst. 236  
Een samenvatting van de belangrijkste resultaten van dit onderzoek verscheen in The Lancet 1986; II: 282.
23. Hemoglobine en hematocriet als risicofactoren voor myocardinfarct, onderzocht in een bedrijfsgeneeskundige setting. 250  
aangeboden voor publicatie.

**Deel VI : Beschouwing en samenvatting.**

24. Beschouwing. 268
25. Samenvatting. 281
- Summary 286
- Dankwoord 291
- Curriculum Vitae 293

Inleiding.

## 1. Motivering.

Diagnostische gegevens en de wijze waarop zij geïnterpreteerd worden zijn van groot belang voor wat er met patiënten in de gezondheidszorg gebeurt. Mensen komen in contact met een arts, al dan niet vanwege bepaalde klachten, en er vindt, al of niet op indicatie, diagnostiek plaats. Op grond daarvan worden mensen "ingedeeld" in bepaalde categorieën, zoals 'afwijkend' of 'normaal', 'ziekte A' en/of 'ziekte B'. Dit leidt weer tot besluitvorming met betrekking tot al of niet, en zo ja hoe, behandelen of vervolgen. In het samenspel tussen de toegenomen bereikbaarheid van medische zorg en de snelle medisch-technologische ontwikkeling is in dit verband de, inmiddels al haast klassieke, vraag op tafel gekomen: hoe kunnen artsen en patiënten doorzeilen tussen de Scylla van het missen van belangrijke diagnosen en behandelingsmogelijkheden en de Charybdis van de medicalisering en de iatrogenese? (1, 2).

Hierbij doen zich onder meer de volgende vragen voor:

- 1) Hoe informatief zijn anamnesticke en andere diagnostische gegevens, om de patiënt op grond hiervan in de juiste 'categorie' in te delen?
- 2) Indien men dit niet met zekerheid kan doen, welke kans is er dan op een juiste respectievelijk verkeerde diagnostische conclusie, als men een bepaalde keuze maakt?
- 3) Wat is de prognostische betekenis van diagnostische informatie?
- 4) Is de patiënt gebaat met de resultaten van het diagnostisch proces, in die zin dat zij van belang zijn voor het therapeutische beleid en de beïnvloeding van de prognose?

In dit proefschrift komen, mede aan de hand van voorbeelden, vooral de vragen 1) tot en met 3) aan de orde. De vierde vraag komt aan bod in het kader van het besliskundige gedeelte.

De genoemde vragen bevatten veel aanknopingspunten voor de klinisch-epidemiologische benadering. In deze benadering staan de klinische (diagnostische, therapeutische en prognostische) aspecten van bepaalde ziekten of gezondheidsproblemen centraal, met gebruikmaking van in-

zichten en onderzoeksmethoden uit de epidemiologie. 'Klinisch' slaat daarbij niet alleen op de kliniek in de zin van 'ziekenhuis', maar betreft ook het veld van de extramurale gezondheidszorg. In dit proefschrift is de klinisch-epidemiologische benadering op verschillende manieren, theoretische en praktische, toegepast. Veel aandacht is besteed aan het belang van de verschillen tussen de situatie van de huisarts en die van de specialist. Immers, vanaf de huisarts naar de specialist vindt een zodanige selectie van patiënten plaats dat communicatiestoornissen tussen beider disciplines over het belang van diagnostische gegevens wel haast onvermijdelijk zijn. Verheldering van de mechanismen en de consequenties van deze selectie is daarom nodig. Voor de eerstelijns geneeskunde is daarbij, naast het optimaliseren van de patiëntenzorg, ook de ontwikkeling van de eigen identiteit in het geding. Deze laatste opmerking wil ik hier aankleden met een korte balans van de positie van de huisarts.

## 2. Speciale aandacht voor de positie van de huisarts.

- 1) Zoals het rapport 'Hoe helpt de dokter' van het Nederlands Huisartsen Genootschap in 1975 stelde, is de huisarts te localiseren op het kruispunt van twee systemen: de geneeskunde en de eerstelijns hulpverlening (3). Lamberts sprak enige jaren later over de huisartsgeneeskunde 'als leer der kruispunten in de gezondheidszorg' (4). Het staan op een kruispunt betekent voor de huisarts een uitdaging, en wel om de juiste richtingen aan te geven. Maar het is een moeilijke, en niet zelden onbevredigende positie. Beide systemen worden steeds ingewikkelder, en minder goed te overzien. Er zijn dwaalsporen te over. Om nog maar te zwijgen van de keren dat de huisarts niet zozeer het kruispunt is van twee systemen, maar de plek waar beide met elkaar in botsing komen.
- 2) Een belangrijk dilemma is dat de huisarts, voorzien van beperkte middelen, met een grote dosis onzekerheid móet kunnen werken. Streeft hij voortdurend naar zekerheid, zoals de specialist dit naar aanleiding van gerichte vraagstellingen meer zal doen, dan brengt dit een groot risico aan overbelasting en beschadiging van de patiënt met zich mee.
- 3) In medische kringen heeft de huisartsgeneeskunde, in haar genera-

lisme wellicht het moeilijkste curatieve specialisme, bij lange na nog niet de waardering gekregen die zij verdient. Dat neemt niet weg dat er steeds meer specialisten te vinden zijn, die inzicht hebben in de specifieke situatie en problematiek van de huisarts.

Een andere kring, namelijk die van de Medische Tuchtcolleges, heeft hier soms nog moeite mee. De ophef rond de uitspraak met betrekking tot de beoordeling van een geval van - uiteindelijk - borstkanker heeft dit onderstreept. Zelfs de kundigheid van huisartsen om een vrouwenborst te onderzoeken werd door het tuchtcollege in twijfel getrokken (5).

4) Aan een duidelijker invulling van de specifieke deskundigheid van de huisarts bestaat in het licht van het voorgaande grote behoefte. Hierover bestaat inhoudelijk echter nog geen overeenstemming onder beroepsgenoten. Weliswaar wordt er veel ontwikkelingswerk verricht, maar de rapporten over de taak- en functie-omschrijving van de huisarts dragen nog een globaal en vrijblijvend karakter (6, 7). In dit verband is een tendens te bespeuren om terug te keren van al te pretentieuze interpretaties van het begrip integrale zorg naar het tenminste goed beheersen van het 'medisch ambacht' en de hierin besloten medische besluitvorming. Hiertoe wordt onder andere gewerkt aan het ontwikkelen van huisartsgeneeskundige strategieën en protocollen (8, 9). Een grote moeilijkheid is dat voor de problemen en klachten waarmee de huisarts te maken krijgt, de aanpak zoals die tijdens de medische opleiding - grotendeels gedragen door specialistische kennis, ervaring en traditie - wordt onderwezen, onvoldoende houvast biedt. Er dient een eigen huisartsgeneeskundig kennisgebied ontwikkeld te worden.

5) Dit alles gaat gepaard met een crisis van een belangrijke achterban van de huisartsgeneeskunde: de geneeskunde zelf. Wat zijn de opbrengst en de kosten ervan in termen van gezondheidswinst en levensgeluk? Ten aanzien van deze kosten komt steeds meer onderzoek van de grond. Gezien de eerste resultaten niet ten onrechte. Zo vonden Steel c.s. bij onderzoek op een interne afdeling van een Amerikaans universiteitsziekenhuis dat van 815 patiënten 36% één of meer iatrogene aandoeningen kreeg (er waren in totaal 497 van dergelijke aandoeningen). Bij 9% waren deze levensbedreigend of invaliderend, en bij 2% van de patiënten droegen ze bij tot het

overlijden. Geneesmiddelen waren verantwoordelijk voor ruim 200 complicaties (10).

Het probleem van de verwijzende huisarts is niet alleen meer: hoe breng ik de mensen de heilzame mogelijkheden van de moderne geneeskunde, maar ook: hoe voorkom ik dat zij aan de gevaren ervan worden blootgesteld?

- 6) Na de vorige punten is het duidelijk dat wetenschappelijk onderzoek in en vanuit de huisartsgeneeskunde erg belangrijk is. De ontwikkeling ervan gaat echter minder snel dan men zou willen. Zowel de specifieke methodologische problemen die erbij aan de orde zijn, als het ontbreken van de nodige infrastructuur, zijn hier debet aan (11). Wanneer het niettemin mogelijk blijkt om, op grond van onderzoek, algemeen gehanteerde medische kennis-elementen vanuit eerstelijns-optiek te herzien, is te verwachten dat verworvenheden als methodisch werken (12), preventie van somatische fixatie (1, 13) en gezins- en levensloopgeneeskunde (14) meer dan tot nu toe tot hun recht zullen komen.

### 3. Opzet van het proefschrift.

In de regel beginnen diagnostische activiteiten naar aanleiding van klachten. Aan het verband tussen klachten en diagnostische bevindingen is in het eerste gedeelte van dit proefschrift dan ook veel aandacht besteed. Eerst wordt aan de hand van basisbegrippen uit de klinische epidemiologie een analyse gegeven van de wijze waarop diagnostische gegevens geïnterpreteerd kunnen worden. Daarna wordt aandacht gevraagd voor de vertekening die kan optreden als men het verband tussen klachten en diagnostische bevindingen bestudeert in populaties die geselecteerd zijn door hulpvraaggedrag, diagnostische activiteiten en/of verwijzing.

Vervolgens worden in deel II een aantal veldstudies beschreven. Daarbij is de vraag gesteld, in hoeverre de vaak gedoceerde verbanden tussen de veelvoorkomende klachten hoofdpijn en moeheid enerzijds, en hoge bloeddruk respectievelijk bloedarmoede en ijzergebrek anderzijds op dit moment in onze streken relevant zijn. Deze studies zijn verricht in verschillende settings, te weten de algemene bevolking, het spreekuur van de huisarts, een voor huisartsen toegankelijk diagnos-



tisch centrum, een polikliniek verloskunde en enkele eerstelijns verloskundigenpraktijken. Daarna wordt wat betreft enkele klachten (moeheid en pijn op de borst) een globale beschrijving gegeven van de diagnostische eindconclusies van huisartsen respectievelijk hun praktisch handelen, mede om een indruk te geven van de mate waarin selectie door middel van verwijzing plaatsvindt.

In het derde gedeelte komt de medisch besliskundige benadering aan de orde, en enkele voorbeelden hiervan worden uitgewerkt. Ook hier wordt aandacht besteed aan verschillen tussen de situaties van huisartsen en specialisten, wat betreft prevalenties, subjectieve kansschattingen en aan te bevelen strategieën. De nadruk wordt gelegd op de verheldering en de illustratie van algemene concepten. Inzicht in de algemene concepten van de besliskunde is naar mijn mening voor de ontwikkeling van het medisch werk van de diverse disciplines vooralsnog belangrijker dan gedetailleerde cijfermatige analyses.

Omdat in de diagnostiek bij voortduring getracht wordt onderscheid te maken tussen 'normale' en 'afwijkende' uitslagen, wordt in het vierde gedeelte geïnventariseerd welke betekenissen men kan toekennen aan het begrip 'normaal', en tot welke verschillende wijzen van interpretatie van diagnostische gegevens dit kan leiden. Hierop aansluitend wordt een analyse gegeven van de invloed van selectie door middel van verwijzing op het onderscheidend vermogen van diagnostische tests.

In deel V wordt verslag gedaan van twee onderzoeken betreffende de prognostische betekenis van een diagnostische bevinding. Het gaat hierbij om het verband tussen:

- het hemoglobinegehalte respectievelijk de hematocriet van zwangeren en de kans op partus prematurus respectievelijk het krijgen van een kind met een laag geboortegewicht;
- het hebben van een (matig) hoog hemoglobinegehalte en de kans op het krijgen van een hartinfarct.

In het zesde en laatste gedeelte wordt het gepresenteerde materiaal geëvalueerd, en worden aanbevelingen gedaan met betrekking tot het klinisch-epidemiologisch onderzoek op het gebied van diagnostiek, met bijzondere aandacht voor de huisartsgeneeskunde.

In dit proefschrift is materiaal opgenomen dat reeds in een bepaalde vorm is of zal worden gepubliceerd. Dat heeft als voordeel, dat de meeste hoofdstukken als afgerond verhaal zijn te lezen. Daartegenover staat het nadeel dat op enkele plaatsen herhalingen voorkomen. Wij vragen de lezer om dit eerder als een tegemoetkoming te beschouwen dan als hinder: de betreffende passages kunnen wat sneller worden doorgenomen, dan wel als zelf-evaluatie materiaal worden opgevat.

#### LITERATUUR.

1. Huygen FJA. Preventie van somatische fixatie. Huisarts en Wetenschap 1978; 21: 363-365.
2. Knottnerus JA. Principes van besliskunde. Nieuw Kompas voor de huisarts, 1984, II/4-15.
3. Hoe helpt de dokter? Nederlands Huisartsen Genootschap, Utrecht 1975.
4. Lamberts H. De huisartsgeneeskunde als leer der kruispunten in de gezondheidszorg. Huisarts en Wetenschap 1978; 21: 465-476.
5. Centraal Medisch Tuchtcollege. Huisarts berispt na late diagnose mammacarcinoom. Medisch Contact 1982; 37: 250.
6. Commissie Takenpakket Landelijke Huisartsen Vereniging. De taken van de huisarts. Medisch Contact 1977; 32: 765.
7. Basistakenpakket van de huisarts, Landelijke Huisartsen Vereniging, 1983.
8. Grol R, Mesker P (red.). Huisarts en toetsing. Nijmeegs Universitair Huisartsen Instituut 1984.
9. Sips ABJI. Protocollen voor de huisarts. Medisch Contact 1986; 41: 277-279.
10. Steel K. Iatrogenic illness on a general medical service at a university hospital. New Engl J Med 1981; 304: 638-642.
11. Touw-Otten F. Wetenschapsbeoefening en huisartsgeneeskunde, Van Loghem Slaterus, Deventer 1981.
12. Holten-Vriesema J, Tompot C, van Aalderen H et al. Methodisch Werken, over een algemene methode van hulpverlening en de opbouw van een functionele relatie, toegespitst op de huisartspraktijk. Huisarts en Wetenschap 1978; 21: 322-335.
13. Grol R (red.). Huisarts en somatische fixatie. Nijmeegs Universitair Huisartsen Instituut, Nijmegen 1981.
14. Huygen FJA. Family Medicine, the medical life history of families, Dekker & van de Vegt, Nijmegen 1978.

1. 關於... 2. 關於... 3. 關於... 4. 關於... 5. 關於... 6. 關於... 7. 關於... 8. 關於... 9. 關於... 10. 關於... 11. 關於... 12. 關於... 13. 關於... 14. 關於... 15. 關於... 16. 關於... 17. 關於... 18. 關於... 19. 關於... 20. 關於... 21. 關於... 22. 關於... 23. 關於... 24. 關於... 25. 關於... 26. 關於... 27. 關於... 28. 關於... 29. 關於... 30. 關於... 31. 關於... 32. 關於... 33. 關於... 34. 關於... 35. 關於... 36. 關於... 37. 關於... 38. 關於... 39. 關於... 40. 關於... 41. 關於... 42. 關於... 43. 關於... 44. 關於... 45. 關於... 46. 關於... 47. 關於... 48. 關於... 49. 關於... 50. 關於... 51. 關於... 52. 關於... 53. 關於... 54. 關於... 55. 關於... 56. 關於... 57. 關於... 58. 關於... 59. 關於... 60. 關於... 61. 關於... 62. 關於... 63. 關於... 64. 關於... 65. 關於... 66. 關於... 67. 關於... 68. 關於... 69. 關於... 70. 關於... 71. 關於... 72. 關於... 73. 關於... 74. 關於... 75. 關於... 76. 關於... 77. 關於... 78. 關於... 79. 關於... 80. 關於... 81. 關於... 82. 關於... 83. 關於... 84. 關於... 85. 關於... 86. 關於... 87. 關於... 88. 關於... 89. 關於... 90. 關於... 91. 關於... 92. 關於... 93. 關於... 94. 關於... 95. 關於... 96. 關於... 97. 關於... 98. 關於... 99. 關於... 100. 關於...

## DEEL I

### INTERPRETATIE VAN DIAGNOSTISCHE GEGEVENS EN DE VERSCHILLEN TUSSEN EERSTE- EN TWEEDELIJNSGENEESKUNDE.

# Interpretatie van diagnostische gegevens

Een onontgonnen terrein in de huisartsgeneeskunde

J. A. KNOTTNERUS\*

Het is al vaak gezegd: de huisarts heeft een werkterrein dat niet door de specialist bestreken kan worden. Huisartsgeneeskunde is een geheel eigen vakgebied, en huisartsgeneeskundige ervaringen en onderzoeken zijn onmisbaar voor het ontwikkelen van een eigen hulpverleningsstrategie. Dit geldt ook voor de interpretatie van diagnostische gegevens. In dit artikel wordt aan de hand van voorbeelden duidelijk gemaakt, dat hier een terrein braak ligt.

## Inleiding

Onlangs vestigden Van der Velden en De Melker de aandacht op de betekenis van de frequentie waarmee aandoeningen voorkomen, voor de interpretatie van diagnostische gegevens. Daarbij werd er op gewezen dat huisarts en specialist te maken hebben met verschillende morbiditeitspatronen, en op grond daarvan ook andere verwachtingen koesteren ten aanzien van eventuele pathologie.<sup>1,2</sup>

Dit artikel sluit hierop aan. Achtereenvolgens zullen worden besproken:

- basisbegrippen;
- de verandering van de voorspellende woorden van een testuitslag bij verwijzing;
- de diagnostische betekenis van klachten en symptomen;
- de mogelijke vertekening van de samenhang tussen symptomen en ziekten ten gevolge van het verwijzingsproces;
- de caesuur tussen „normale” en „afwijkende” uitslagen;
- de invloed van de mate van ontwikkeling van het klinisch beeld op testuitslagen.

## Basisbegrippen

We gaan uit van vrouwen in de leeftijd van 20 tot 30 jaar, bij wie de huisarts door middel van fysische diagnostiek een palpabele afwijking in de borst vindt. De hierbij gebruikte getallen zijn

zoveel mogelijk gebaseerd op de literatuur.

Na het lichamelijk onderzoek zijn er twee conclusies mogelijk:

- suspect voor borstkanker (positieve uitslag);
- niet suspect voor borstkanker (negatieve uitslag).

In *figuur 1* is weergegeven hoe de onderzoeksresultaten kunnen uitvallen bij 1000 vrouwen. Er is sprake van 28 terecht-positieve (TP), 38 fout-positieve (FP), 22 fout-negatieve (FN) en 912 terecht-negatieve (TN) uitslagen.

De huisarts zal nu primair geïnteresseerd zijn in de beantwoording van de volgende vragen:

- Hoe groot is de kans dat vrouwen bij wie ik een suspect knobbeltje vind, inderdaad borstkanker hebben?
- Hoe groot is de kans dat vrouwen bij wie ik het knobbeltje niet suspect vind, geen borstkanker hebben?

De eerste vraag betreft de voorspellende waarde van de positieve testuitslag (VW+), de tweede vraag de voorspellende waarde van de negatieve testuitslag (VW-).

De testmethode zelf wordt gekenmerkt door:

- De kans dat aanwezige ziektegevallen door de test worden ontdekt, de *sensitiviteit*.
- De kans dat personen zonder de ziekte door de test ook als zodanig worden herkend, de *specificiteit*.

De in *figuur 1* vermelde waarden voor sensitiviteit en specificiteit komen overeen met bevindingen in Nederlands en Amerikaans onderzoek.<sup>3,4</sup>

Uit *figuur 1* is voorts de *prevalentie* af te lezen: het percentage personen in de bestudeerde populatie dat de betreffende ziekte werkelijk heeft, in dit geval 5 procent.<sup>1</sup>

## Verandering van voorspellende waarde bij verwijzing

Stel dat de huisarts als beleid heeft:

- elke vrouw met een suspect knobbeltje verwijs ik rechtstreeks naar de chirurg voor biopsie;
- de overige vrouwen met een palpabele afwijking krijgen een mammografie: als die positief uitvalt (ruimte voor verdenking openlaat), verwijs ik alsnog voor biopsie.

Als we er hierbij van uitgaan dat op indicatie verrichte mammografie een sensitiviteit van 91 procent en een specificiteit van 84 procent heeft<sup>4,7</sup>, en dat biopsie uiteindelijk de werkelijke toestand onthult, dan is gemakkelijk na te rekenen wat te verwachten is. In *figuur 2* is dat gedaan.

Het blijkt dat bij de chirurg de prevalen-

*Figuur 1. Mogelijke onderzoeksresultaten van fysische diagnostiek bij 1000 vrouwen met klachten over een knobbeltje in de borst. Basisbegrippen.*

	Borst CA	Geen borst CA	Totaal
Onderzoek positief: suspect	28 TP	38 FP	66
Onderzoek negatief: niet-suspect	22 FN	912 TN	934
Totaal	50	950	1000

**Voorspellende waarde van een positief onderzoeksresultaat:**

$$VW+ = \frac{TP}{TP + FP} = 42\%$$

**Voorspellende waarde van een negatief onderzoeksresultaat:**

$$VW- = \frac{TN}{TN + FN} = 98\%$$

**Sensitiviteit van het onderzoek:**

$$\frac{TP}{TP + FN} = 56\%$$

**Specificiteit van het onderzoek:**

$$\frac{TN}{TN + FP} = 96\%$$

**Prevalentie:**

$$\frac{TP + FN}{TP + FN + FP + TN} = 5\%$$

\* Capaciteitsgroep Gezondheidszorgonderzoek/Epidemiologie van de Rijksuniversiteit Limburg, Maastricht.

tie van borstkanker 21 procent wordt. Als de chirurg nu, alvorens tot biopsie over te gaan, palpatie verricht (en dat niet beter of slechter doet dan de huisarts dan heeft deze palpatie een positief voorspellende waarde (VW+) van 79 procent en een negatief voorspellende waarde (VW-) van 89 procent. De chirurg zal dus bij palpatie naar verhouding minder fout-positieven vinden dan de huisarts, en ten onrechte kan de indruk ontstaan dat hij dus ook beter palpeert. Ook al zou hij relatief vaker een fout-negatieve diagnose stellen, dan nog valt dit minder op, daar hij dikwijls ook een biopsie zal doen. Bovendien zal hij de meeste vrouwen bij wie geen biopsie wordt gedaan, na korte tijd zelf nog eens controleren, zodat hij fout-negatieven in tweede instantie toch zal vinden.

De (weinig) fout-negatieven van de diagnostische procedure van de huisarts - dus degenen die ten onrechte niet verwezen zijn na een niet verdachte mammografie-uitslag - zullen voorlopig gerustgesteld zijn en wellicht ook niet allemaal voor controle komen opdagen. Dergelijke gevallen worden later als „gemiste“ carcinomen beschouwd, dan wel aangemerkt als voorbeelden van *doctors delay*.

Het is duidelijk dat het missen van een aantal gevallen in eerste instantie alleen voorkomen had kunnen worden, als de huisarts bij alle vrouwen met een palpabele afwijking biopsie had laten verrichten. Dat had in dit geval betekend dat er 768 extra biopsieën gedaan hadden moeten worden om 2 extra gevallen al in eerste instantie te ontdekken.

Men kan inzien dat bij de keuze tussen de diverse diagnostische strategieën (alleen verwijzing bij suspecte palpatie, de beschreven strategie, of: iedere vrouw met klachten of een knobbelte verwijzen) van belang is met welke prevalenties men te maken heeft. Als hiervan een goede indruk bestaat, kan men voor elke strategie een schatting maken van de kans op gemiste carcinomen en de kans op biopsieën bij gezonde vrouwen.<sup>4</sup> Relevante gegevens daarbij zijn de wetenschap dat een knobbelte bij ouderen veel vaker op kanker duidt dan bij jongeren,<sup>1</sup> en het feit dat kans op kanker bij mammaklachten in het algemeen veel kleiner is dan bij de aanwezigheid van een knobbelte.

#### Voorselectie door de huisarts

Het bovenstaande resumerend kan gesteld worden: Wanneer de prevalentie toeneemt, neemt de voorspellende

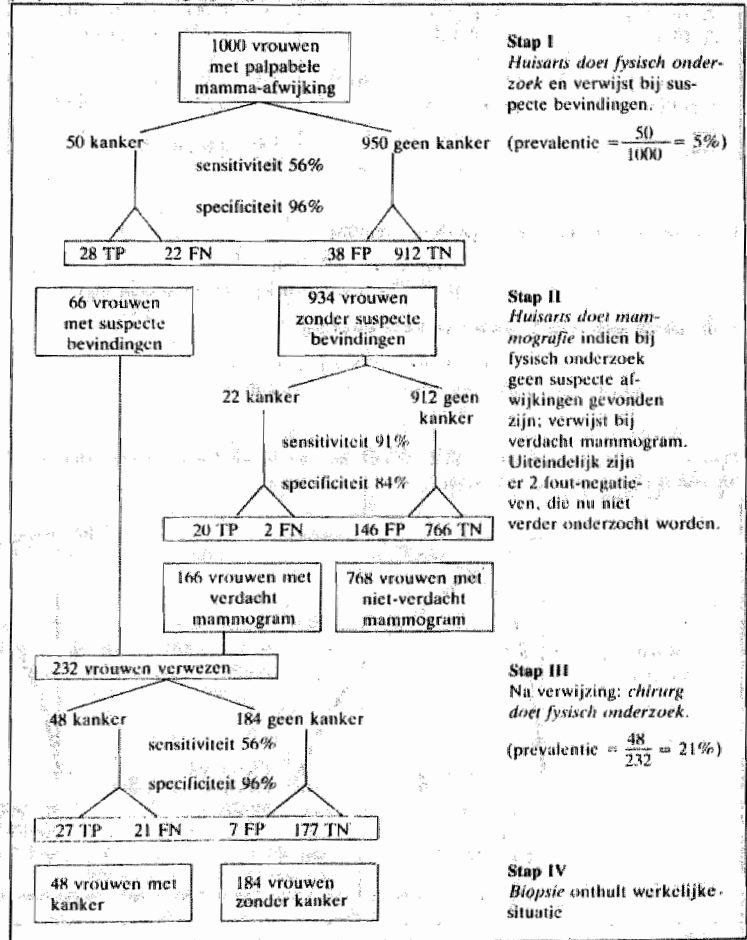
waarde van een positieve testuitslag toe, maar wordt de voorspellende waarde van een negatieve testuitslag kleiner: er komen minder fout-positieve doch meer fout-negatieve uitslagen. Wanneer de prevalentie afneemt, is het omgekeerde te verwachten: de voorspellende waarde van een positieve testuitslag wordt kleiner, die van een negatieve uitslag groter; er zijn meer fout-positieve maar minder fout-negatieve uitslagen te verwachten.

Het feit dat er een positieve samenhang is tussen de voorspellende waarde van een positieve testuitslag en de prevalentie van de aandoening waarop men onderzoekt, zal niet moeilijk te begrijpen zijn. Lastiger is het zich voor te stellen dat de voorspellende waarde van een negatieve uitslag kleiner wordt als de prevalentie toeneemt. Wellicht kan nog

een voorbeeld verhelderend werken. Bij een vrouw van 25 jaar zonder klachten wordt bij een routinecervixuitstrijkje een Pap I gevonden. De arts zal er dan op vertrouwen dat er geen sprake is van cervixkanker. Stel nu dan de uitslag Pap I ook gegeven wordt bij een 40-jarige vrouw met klachten van tussentijds bloedverlies en contactbloedingen. Ondanks het feit dat de testuitslag ook nu negatief is, zal de arts in dit geval minder zeker van zijn zaak zijn: hij houdt rekening met een fout-negatieve uitslag, en acht kanker nog geenszins uitgesloten: de VW- wordt duidelijk lager geacht.

De arts past nu in feite toe wat zojuist geformuleerd is: hij interpreteert de testuitslag met in het achterhoofd de kennis dat cervixkanker bij vrouwen van 25 jaar zonder klachten zeer weinig

Figuur 2. Mogelijke resultaten van diagnostische procedure bij jonge vrouwen met palpabele mamma-afwijking.



voorkomt (de prevalentie in zo'n populatie is erg laag), en dat de ziekte bij vrouwen van 40 jaar met de beschreven klachten veel vaker voorkomt (de prevalentie in een dergelijk omschreven populatie is hoger).

Als we deze principes vervolgens toepassen op de verschillende situaties waarin de huisarts en de specialist verkeren, zijn de volgende conclusies te trekken, (waarbij we ervan uitgaan dat beiden in beginsel even kundig zijn).

- In de populatie patiënten die de specialist bezoekt, is - als gevolg van de voorselectie door de huisarts - de prevalentie van ernstige ziekten en van ziekten die specialistische behandeling vergen, groter dan in de patiëntenpopulatie die de huisarts ziet. Dat betekent

dat de voorspellende waarde van een afwijkende bevinding bij een door de specialist verricht onderzoek altijd groter is dan dezelfde bevinding door de huisarts. Ten aanzien van de afwezigheid van deze ziekten is de voorspellende waarde van een niet-afwijkende bevinding bij de huisarts juist groter.

- Voor de ziekten die de huisarts vaak ziet, en de specialist minder vaak, geldt het omgekeerde. Voorbeelden hiervan zijn vele *minor ailments*, exanthemen bij kinderen en dergelijke.

- De voorspellende waarde van een positieve uitslag van een specialistisch onderzoek kan relatief groot zijn, doordat de huisarts voorselecteert. Als de huisarts dit niet deed, zou veel overbodige diagnostiek en eventueel behande-

ling plaatsvinden. Men zou kunnen zeggen: doordat de huisarts een grote negatief voorspellende waarde heeft, kan de specialist een hoge positief voorspellende waarde halen.

Verder speelt ook nog een rol dat de huisarts door middel van anamnese en lichamelijk onderzoek die patiënten selecteert bij wie hij een verhoogde kans op bepaalde aandoeningen aanwezig acht. Daarmee verhoogt de huisarts ook de voorspellende waarde van afwijkende bevindingen bij nader onderzoek dat in eigen beheer plaatsvindt.<sup>2</sup>

In *figuur 3* is het beschreven verband tussen sensitiviteit, specificiteit, prevalentie en voorspellende waarden samengevat in de formule van Bayes. Deze formule is overigens direct af te leiden uit de vierveldentabel en geeft in principe ook niet meer informatie dan deze tabel.

In *figuur 4* zijn de verschillende verbanden grafisch weergegeven voor een tweetal onderzoeken, palpatie en mammografie.

### Betekenis van klachten en symptomen

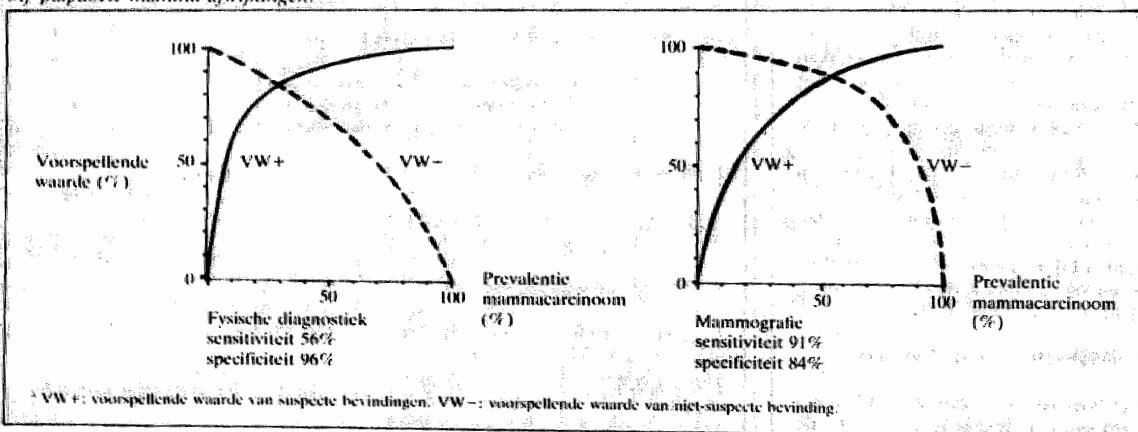
Niet alleen onderzoeksresultaten, maar ook symptomen en klachten op zich zelf hebben een voorspellende waarde. Zij zijn in die zin niets anders dan een „test“. In het eerste voorbeeld spraken we van de prevalentie van borstkanker in een groep vrouwen met palpabele afwijkingen. We hadden ook kunnen spreken van de voorspellende waarde van een palpabele afwijking voor het hebben van borstkanker. Ook

*Figuur 3. De formule van Bayes.*

	Ziek	Niet-ziek
Testuitslag positief	TP = sensitiviteit x prevalentie	FP = (100%-specificiteit) x (100%-prevalentie)
Testuitslag negatief	FN = (100%-sensitiviteit) x prevalentie	TN = specificiteit x (100%-prevalentie)
Totaal	prevalentie	100%-prevalentie

- 1) voorspellende waarde van een positieve testuitslag =  $\frac{TP}{TP + FP}$
- =  $\frac{\text{sensitiviteit} \times \text{prevalentie}}{(\text{sensitiviteit} \times \text{prevalentie}) + (100\% - \text{specificiteit}) \times (100\% - \text{prevalentie})}$
- 2) de voorspellende waarde van een negatieve testuitslag =  $\frac{TN}{TN + FN}$
- =  $\frac{\text{specificiteit} \times (100\% - \text{prevalentie})}{\text{specificiteit} \times (100\% - \text{prevalentie}) + (100\% - \text{sensitiviteit}) \times \text{prevalentie}}$

*Figuur 4. Verband tussen sensitiviteit, specificiteit en voorspellende waarden van fysische diagnostiek, respectievelijk mammografie bij palpabele mamma-afwijkingen.<sup>2</sup>*



deze waarde is gelijk aan 5 procent, zolang er nog niet gedifferentieerd is tussen suspecte en niet suspecte afwijkingen (c.q. de „tweede test“ is uitgevoerd).

Aan de hand van enkele voorbeelden zal nader worden ingegaan op de betekenis van klachten en symptomen.

**Appendicitis acuta en pijn rechts in de onderbuik.** Uit morbiditeitsgegevens mogen we concluderen dat appendicitis acuta zich per jaar gemiddeld 4 à 5 keer voordoet in een praktijk van 2500 zielen.<sup>8-10</sup> Bij klinisch onderzoek bleek circa 75 procent van de patiënten met appendicitis acuta pijn rechts in de onderbuik te hebben.<sup>11</sup> De sensitiviteit van deze klacht is dus 75 procent ten aanzien van appendicitis.

Als we nu, mede op grond van klachtenregistraties, schatten dat zich in een praktijk van genoemde grootte per jaar ongeveer 40 mensen melden met pijn rechts in de onderbuik, kunnen we de VW+ van de klacht berekenen. Van de 4 mensen met appendicitis krijgen er 3 (75 procent) pijn rechts in de onderbuik. Totaal melden zich echter 40 personen met die klacht. Aangenomen dat de drie genoemde patiënten allen de hulp van de huisarts inroepen, is de VW+ van de klacht slechts  $3/40 = 7,5$  procent.

Pijn rechts in de onderbuik op zichzelf lijkt hier dus van betrekkelijke waarde bij het voorspellen van de aanwezigheid van appendicitis. Die waarde zou nog geringer worden als we rekening houden met de mogelijkheid dat de sensitiviteit van de klacht voor de ziekte in de huisartspraktijk geringer is, doordat het beeld wat minder ver ontwikkeld zal zijn dan in de kliniek.

**Bultje in de hals.** Wanneer één bepaald symptoom kan wijzen op verschillende aandoeningen, is het van belang dat men zich realiseert met welke populatie men te maken heeft. Met de populatie kan immers ook de voorspellende waarde van dat symptoom variëren. We geven weer een voorbeeld.

Bij de arts meldt zich een patiënte met een bultje in de hals, dat naar haar zeggen reeds enkele weken aanwezig is. Het is licht drukgevoelig, glad en 0,5 cm in doorsnede.

Het is duidelijk dat een dergelijke presentatie bij de huisarts een geheel andere betekenis heeft dan op het spreekuur van de oncoloog. Op grond van exact dezelfde bevindingen zal de eerste een ontstoken klier het meest waarschijnlijk achten en een afwachende houding

aannemen. De laatste zal echter een primaire maligniteit of metastase willen uitsluiten met behulp van verder onderzoek. Men kan zich ook voorstellen dat, wanneer beiden een klinische les over dit symptoom zouden geven, zij met sterk verschillende aanbevelingen zouden eindigen.

**Hemoglobinegehalte en „anemiekklachten“.** Op grond van onderzoek in de algemene bevolking is het onwaarschijnlijk, dat er in onze streken een samenhang bestaat tussen een laag hemoglobinegehalte en de belangrijkste klassieke „anemiesymptomen“.<sup>12</sup> Huisartsen hebben echter tijdens hun studie geleerd dat anemie bepaalde klachten veroorzaakt, en de kans dat zij het hemoglobinegehalte bepalen, is daarom bij aanwezigheid van die klachten groter dan bij klachten waarvan een verband met anemie niet geboekt staat. De kans dat door toeval een laag Hb wordt gevonden, is bij mensen met „anemiekklachten“ dan ook groter dan bij andere klachten. Als huisartsen dan mensen met een laag Hb waarvoor zij geen verklaring of oplossing hebben, naar de specialist verwijzen voor nader onderzoek, zal deze laatste naar verhouding veel patiënten zien met een laag Hb, die „anemie-klachten“ hebben. De sensitiviteit van anemiekklachten ten aanzien van de aanwezigheid van een laag Hb zal dan, ook als er geen verband tussen beide is, vrij hoog kunnen zijn. De specialist wordt aldus bevestigd in zijn opvatting dat beide verschijnselen samenhangen. In *figuur 5* wordt dit effect gedemonstreerd aan de hand van een getallenvoorbeeld.

Het is zeer wel mogelijk dat dit fenomeen van waarnemersvertekening met betrekking tot de samenhang tussen symptomen en beschreven ziektebeelden wijd verbreid is.

### Optimale afkappunten

Het punt waar men de caesuur legt tussen een niet-afwijkende en een afwijkende uitslag, is zelden een objectief gegeven. Meestal gaat het om een arbitrair besluit op grond van consensus. Men kan dan spreken van een vastgesteld „afkappunt“ voor de beslissing tussen normaal en afwijkend. Voorbeelden zijn: de ondergrens van het Hb, de bovengrens van de bloeddruk, leverfuncties of de bloedsuiker; de hoogte van het ST-segment op het inspannings-ECG en de hoeveelheid acceptabele mitosen bij pathologisch-anatomisch onderzoek.

De sensitiviteit en de specificiteit van het onderzoek zijn afhankelijk van de keuze van het afkappunt. In *figuur 6a* wordt dit duidelijk aan de hand van de verdelingen van hematocrietwaarden (Ht) bij Zweedse vrouwen met en zonder ijzergebrek. Garby kon deze onderverdeling maken, afgaande op de respons op ijzertoeiening bij een steekproef van vrouwen uit de algemene bevolking.<sup>13</sup>

Te zien is dat, als het afkappunt voor ijzergebreksanemie wordt verhoogd (dus minder stringent wordt), meer ijzergebreksanemie-patiënten worden opgespoord, terwijl er tegelijkertijd

*Figuur 5. Hypothetisch rekenvoorbeeld van de gesuggereerde samenhang van „anemiekklachten“ en een laag Hb (mmol/l), ten gevolge van selectie tijdens het medisch hulpverleningsproces.*

1. In de praktijkpopulatie van de huisarts is de situatie als volgt:  
*Anemiekklachten en Hb bij 1000 patiënten.*

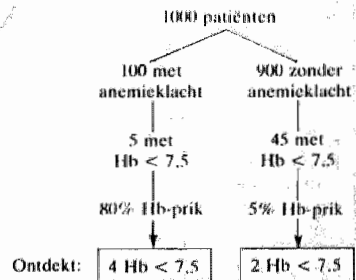
Anemie-klachten	Hb < 7,5	Hb ≥ 7,5	Totaal
+	5	95	100
-	45	855	900
Totaal	50	950	1000

Er is geen verband tussen het Hb en de anemiekklachten; het percentage mensen met anemiekklachten is bij mensen met Hb < 7,5 even groot als bij mensen met Hb ≥ 7,5:

$$5/50 = 95/950 = 10\%$$

2. Bij een bezoek aan de huisarts is de kans op een Hb-prik bij anemiekklachten 80% en bij andere klachten 5%:

*Ontdekking van Hb < 7,5 bij 1000 patiënten.*



3. Bij gelijke kansen op verwijzing voor analyse wegens onbegrepen anemie: percentage mensen met anemiekklachten bij mensen met Hb < 7,5 is bij de specialist  $4/6 = 67\%$



meer fout-positieve meetresultaten zijn: de sensitiviteit neemt toe, doch de specificiteit wordt kleiner. Verlaagt men de grenswaarde (er wordt dan een stringenter criterium gekozen), dan gebeurt precies het omgekeerde. Men kan ook stellen: naarmate men bij onderzoek van een bepaalde populatie een meer stringente grenswaarde kiest, is de voorspellende waarde van een afwijkende testuitslag groter. Bij een minder strenge norm wordt de voorspellende waarde van een afwijkende uitslag kleiner.

De vraag is nu: wat is het optimale afkappunt bij een test?

Deze vraag is niet eenvoudig en zeker niet eenduidig te beantwoorden. Het antwoord is in de eerste plaats afhankelijk van de subjectieve waarde die pa-

tiënt en arts toekennen aan een fout-positieve c.q. fout-negatieve uitslag: als men de consequenties van een fout-negatieve uitslag ernstiger vindt dan die van een fout-positieve, zal men de nadruk leggen op een test met een hoge sensitiviteit. Worden daarentegen fout-positieven ernstiger geacht, dan is een meer specifieke test extra belangrijk. Het is voor de lezer wellicht een goede vingeroefening, eens na te gaan hoe dit wat hem of haar betreft ligt voor verschillende onderzoeken en aandoeningen. Op dit moment lijkt het er op, dat fout-negatieve uitslagen in het algemeen ernstiger worden gevonden dan fout-positieve, ook indien de laatste kunnen leiden tot onnodig, vaak ingrijpend vervolgonderzoek of overbodige therapie. De huisarts zal dan ook vaak

geneigd zijn het afkappunt niet te stringent te stellen. Deze beslissing impliceert echter een keuze voor een groter aantal fout-positieven, met als mogelijk gevolg: meer iatrogenese en somatische fixatie.

De optimale keuze van het afkappunt hangt ook af van de prevalentie van de mogelijke aandoening<sup>14</sup>. Men kan dit zien in *figuur 6b*: wanneer er meer vrouwen met ijzergebreksanemie in de populatie voorkomen, zijn er bij een zelfde afkappunt meer fout-negatieve uitslagen te verwachten. Wil men deze in toom houden, dan zal het afkappunt minder stringent gekozen moeten worden.

Het is om deze reden niet vanzelfsprekend dat in situaties waar in verschillende prevalenties aan de orde zijn – bijvoorbeeld de situatie van de huisarts in vergelijking tot de situatie van de specialist – slechts één afkappunt gehanteerd wordt. In het algemeen gebeurt dit echter wel, omdat op de specifieke situatie afgestemde criteria doorgaans ontbreken.

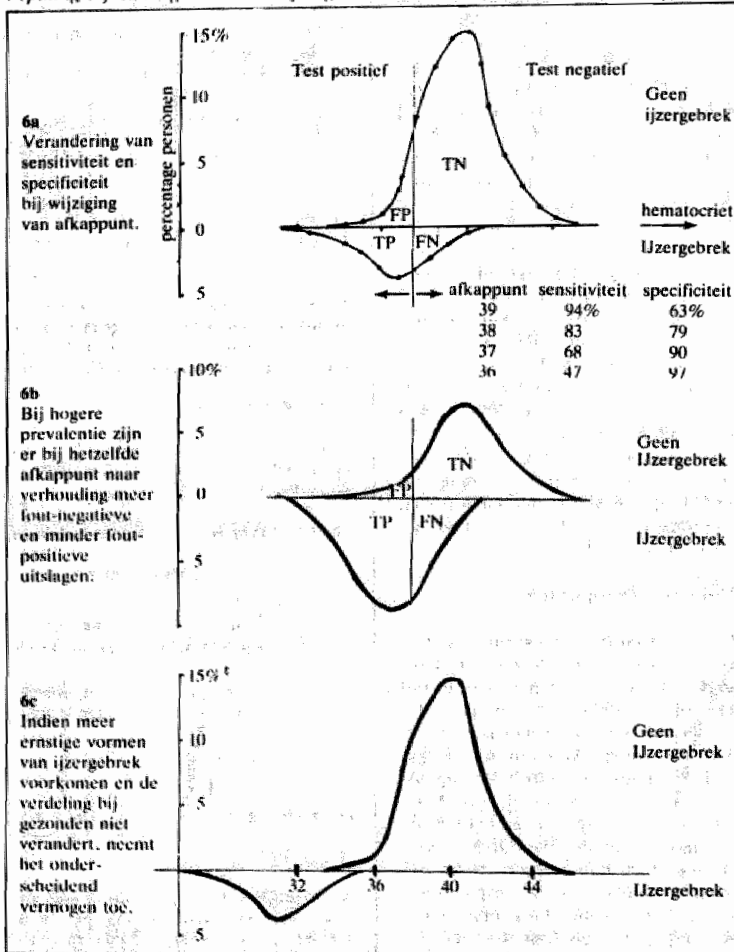
In *figuur 6a* is overigens te zien dat men ook op een andere wijze testuitslagen kan interpreteren dan aan de hand van enkelvoudige afkappunten. Bij iedere Ht-waarde is namelijk de bijbehorende kans op ijzergebrek aan te geven, als quotiënt van de hoogte van de curve bij ijzergebrek en de som van de hoogten van beide curves voor de betreffende Ht-waarde. In het voorbeeld is deze kans bij achtereenvolgens de waarden 31 en lager, 33, 36, 39 en 41 en hoger: 100, 80, 59, 9 en 0 procent. Ook deze kansen zijn echter niet onafhankelijk van de prevalentie.

## De ernst van het klinisch beeld

Een van de belangrijke verschillen in de morbiditeitspatronen, die de huisarts en de specialist onder ogen krijgen, is voorts het ontwikkelingsstadium van het klinisch beeld waarmee aandoeningen zich manifesteren.

In het algemeen zullen de verschijnselen in de fase waarmee de huisarts ermee geconfronteerd wordt, minder duidelijk ontwikkeld zijn dan in de fase na eventuele verwijzing naar de specialist. In de eerste situatie zullen de zieken minder goed te ontdekken zijn en ook minder goed van niet-zieken kunnen worden onderscheiden dan in de tweede situatie. In de huisartsen-fase is het onderscheidend vermogen van onderzoeksmethoden dan ook veelal lager dan in de specialisten-fase. Sensitiviteit en – bij juiste keuze van het afkappunt – ook

**Figuur 6. Relatie tussen afkappunt, sensitiviteit en specificiteit van de hematocrietbepaling bij de diagnostiek van ijzergebrek.**



specificiteit zijn dan voor de specialist altijd gunstiger.

We geven enkele voorbeelden, waarbij we ons beperken tot situaties waarbij in principe een dichotoom onderscheid ziek/niet-ziek te maken is.

Men kan inzien dat, wanneer men te maken heeft met ernstiger gevallen van ijzergebreksanemie, de verdeling van hematocrietwaarden voor „zieken” meer naar links verschoven is (figuur 6c). Het onderscheidend vermogen van de Ht-bepaling ter opsporing van ijzergebrek wordt dan groter, ook al blijft de prevalentie onveranderd. Ook kan men zich voorstellen dat het onderscheidend vermogen van de thermometer in de allereerste fase van appendicitis acuta nihil is, en dat hetzelfde geldt voor leverfunctietests bij beginnende of lichte hepatitis. Teruggrijpend op het voorbeeld over borstkanker: de specialist onderzoekt gemiddeld grotere tumoren dan de huisarts, en deze kunnen beter gevoeld worden.

In feite ligt de zaak nog gecompliceerder, de specialist zal vaker dan de huisarts te maken hebben met meer ontwikkelde en dus beter te palperen stadia van borstkanker: de sensitiviteit is groter. Maar fysiologische en benigne borstafwijkingen zijn door middel van palpatie moeilijk te onderscheiden van kanker<sup>13</sup>. Er worden daardoor ook patiënten verwezen met een suspect beeld die geen kanker hebben, en ook wat deze gevallen betreft, is er een selectie naar de specialist. „Niet-kanker” is dan ook door de specialist niet gemakkelijk door middel van palpatie vast te stellen, en wellicht is de specificiteit van zijn onderzoek zelfs kleiner. Het onderscheidend vermogen van de specialistische palpatie hoeft in dat geval niet per se beter te zijn.

In het algemeen zou men kunnen zeggen: de zieken in de huisartspraktijk lijken meer op gezonden (alleen de herkenbare gevallen worden verwezen) en de gezonden in het ziekenhuis lijken meer op zieken (anders waren zij niet verwezen). Huisarts en specialist hebben dus te maken met een ander spectrum en een andere mate van ontwikkeling van klinische beelden bij zowel zieken als „niet-zieken”. Zij doen in dat opzicht dus ook geheel andere ervaringen op.

Het feit dat de validiteit van onderzoeksmethoden zou moeten worden vastgesteld in een populatie die op relevante punten (zoals de ernst van het klinisch beeld) representatief is voor de populatie waarop de methode zal worden toegepast, staat in de literatuur pas de laatste

jaren in de belangstelling.<sup>16,18</sup> Voor de huisartsgeneeskunde zal dit belangrijke implicaties kunnen hebben, daar weinig methoden „geijkt” zijn in de huisartsensetting.

## Behoeftte aan onderzoek in de huisartspraktijk

De diagnostische posities van huisarts en specialist verschillen uiteraard op meer punten dan in het voorgaande uitgewerkt zijn. De huisarts moet meer dan de specialist werken met nog weinig uitgekristalliseerde vraagstellingen, en de „probleemruimte” is groot.<sup>19,20</sup> Er zijn vaak meer klachten of problemen tegelijk aan de orde in één hulpverlenings-episode, en vele verschillende aandoeningen moeten in de hypothesevorming worden betrokken. De huisarts zal meer dan de specialist afhankelijk zijn van onderzoeksmethoden die een breed veld van mogelijkheden op nog weinig gerichte wijze kunnen „afgrazen”, zoals de hemoglobinebepaling, de meting van de bezinkingssnelheid van de erythrocyten en de meting van de lichaamstemperatuur, en in wezen ook de anamnese en het oriënterend lichamelijk onderzoek. Bovendien zal het vaak niet mogelijk of niet zinvol zijn om tot een definitieve diagnose te komen.

Aan de andere kant heeft de huisarts weer vaak het voordeel, over meer voorkennis over de patiënt te beschikken. Daardoor wordt een betere beoordeling van (extra) risico’s mogelijk. Ook van belang is dat de methode van het afwachten van het natuurlijk beloop een meer gehanteerd en geaccepteerd deel van het diagnostisch repertoire is.

In dit artikel kwam naar voren dat huisarts en specialist door ervaring leren om te gaan met geheel verschillende patronen van voorspellende waarden van klachten, symptomen en onderzoeksresultaten. De tweede leert de werkwijze die bij zijn situatie past, in het algemeen vanaf de aanvang van de studie; de eerste moet later nog veel bijspijkeren, wellicht vaak door schade en schande. De specialist verkeert verder in een situatie waarin hij wordt geconfronteerd met waarnemingen die verbanden kunnen suggereren en bevestigen die er mogelijk niet zijn.

Hoewel men van elkaars ervaringen kan en dient te leren, is het onjuist, en mogelijk zelfs schadelijk, als een van beiden de ander zijn ervaring, kennis en norm als maatgevend opdringt. Zowel het miskennen van pathologie als het bevoor-deren van iatrogenese en somatische

fixaties liggen dan op de loer. De huisarts, die het grootste deel van zijn studie door de specialistische bril heeft leren kijken, wordt met name door het laatste bedreigd.

Voor de huisartsgeneeskunde bestaat de noodzaak, het onontgonnen gebied van de diagnostiek in kaart te brengen. Daarbij moet aandacht worden besteed aan de voorspellende waarde van klachten, symptomen en testuitslagen, aan de besluitvorming over de mate waarin bepaalde uitslagen nog wel en niet meer acceptabel zijn, aan het vinden en verbeteren van de diagnostica die juist voor de huisartspraktijk van belang zijn, en aan het corrigeren van zichzelf bevestigende klinische vertekeningen van de samenhang tussen symptomen en aandoeningen.

<sup>1</sup> Velden HGM van der. Diagnose of prognose. De betekenis van de epidemiologie voor het handelen van de huisarts. Huisarts en Wetenschap 1983; 26:125-8.

<sup>2</sup> De Melker RA. Diagnostische faciliteiten voor de huisarts: wie wordt er beter van? Huisarts en Wetenschap 1983; 26:184-8.

<sup>3</sup> Rombach JJ. Breast cancer screening. Results and implications for diagnostic decision making [Dissertatie]. Alphen a/d Rijn: Stafleu, 1980.

<sup>4</sup> Griner PF, Mayewski RJ, Mushlin AI, Greenland P. Selection and interpretation of diagnostics tests and procedures. Ann Intern Med 1981; 94:533-96.

<sup>5</sup> Clarijs WWJ, Deelen-Willems M van, Hendriks JHCL, Lubbers EJC, Schmidt WJH. Het occulte mammacarcinoom. Ned Tijdschr Geneesk 1977; 121:526-9.

<sup>6</sup> Dronkers DJ. De betrouwbaarheid van mammografie. Ned Tijdschr Geneesk 1977; 123:955-7.

<sup>7</sup> Molenaar JC, Herder-Kroon B den. De praktische waarde van de mammografie voor de chirurg. Ned Tijdschr Geneesk 1970; 114:1688-90.

<sup>8</sup> Oliemans AP. Morbiditeit in de huisartspraktijk [Dissertatie]. Leiden: Stenfert Kroese, 1969.

<sup>9</sup> Lamberts H. De morbiditeitsanalyse van de groepspraktijk Ommoord. Een nieuwe ordening van ziekte- en probleemgedrag in de huisartsgeneeskunde. II. Huisarts en Wetenschap 1975; 18:7-39.

<sup>10</sup> Lamberts H. Incidentie en prevalentie van gezondheidsproblemen in de huisartspraktijk. Huisarts en Wetenschap 1982; 25:401-14.

<sup>11</sup> Dombal FT de. Acute abdominal pain. An O.M.G.E. survey. Scand J Gastroenterol 1979; 14 (suppl 56):29-43.

<sup>12</sup> Elwood PC. Symptoms and circulating haemoglobin level. J Chronic Dis 1969; 21:615-28.

<sup>13</sup> Garby L. The normal haemoglobin level. Br J Haematol 1970; 19:429-34.

- <sup>14</sup> Weinstein MC, Fineberg HV. Clinical decision analysis. Philadelphia: Saunders, 1980.
- <sup>15</sup> Love SM, Gelman RS, Silen W. Fibrocystic "disease" of the breast - a non-disease? *New Engl J Med* 1982; 307:1010-4.
- <sup>16</sup> Ranschoff DF, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. *New Engl J Med* 1978; 299:926-30.
- <sup>17</sup> Anonymous. Help with diagnostic decisions [Editorial]. *Ann Int Med* 1982; 97:444-5.
- <sup>18</sup> Rümke CL. Kanttekeningen over de gevoeligheid, de specificiteit en de voorspellende waarden van diagnostische tests. *Ned Tijdschr Geneesk* 1983; 127:556-61.
- <sup>19</sup> Kooy S van der. De huisarts als hulpverlener bij lichamelijke aandoeningen. *Huisarts en Wetenschap* 1981; 24:454-8.
- <sup>20</sup> Weel C van. Achilles tot hardloper bekeerd. *Huisarts en Wetenschap* 1981; 24:417-23.

# Symptomatologie en selectiebias

Vertekening van het verband tussen klachten en diagnoses ten gevolge van selectie naar hogere echelons

J. A. KNOTTNERUS, P. G. KNIPSCHILD EN F. STURMANS\*

Waarnemingen ten aanzien van het verband tussen klachten of symptomen enerzijds, en ziekten of diagnostische conclusies anderzijds worden sterk beïnvloed door selectieprocessen. Deze selectie vindt plaats vanaf het niveau van de algemene bevolking, via hulpvraag-gedrag, indicatiestellingen door de huisarts, behandelingsresultaten en eventuele verwijzing. Door selectie kunnen verbanden worden gesuggereerd en gereproduceerd die in werkelijkheid niet bestaan, louter als produkt van aangeleerde indicatiestellingen. Het corrigeren van dergelijke vertekeningen door middel van wetenschappelijk onderzoek vanuit de huisartspraktijk is alleen mogelijk indien men de juiste referentiegroep kiest.

## Inleiding

In het klinisch epidemiologisch onderzoek is het trekken van verkeerde conclusies ten gevolge van zogeheten *selectiebias* een belangrijke valkuil. Selectiebias is een vertekening van onderzoeksresultaten die veroorzaakt wordt door de wijze waarop de onderzoeksgroepen worden geselecteerd. *Berkson* vestigde als eerste de aandacht op de vertekening die kan optreden bij onderzoek op basis van ziekenhuispatiënten.<sup>1</sup> Sindsdien is *Berkson's fallacy* (*admission rate bias*) steeds een actueel studieonderwerp gebleven (kader op pagina 326).<sup>2,6</sup>

Met dergelijke selectiebias moet ook rekening worden gehouden bij mogelijke verbanden tussen klachten of symptomen enerzijds, en diagnoses of diagnostische bevindingen anderzijds. Het handelen van de huisarts en verwijzing naar de specialist kunnen schijnverbanden introduceren, en verwachtingen ten aanzien van verbanden tussen klachten en diagnoses kunnen zichzelf waarmaken. Verbanden kunnen zichzelf reproduceren, niet doordat ze werkelijk bestaan, maar doordat men – ten onrechte – geleerd heeft dat ze bestaan. Het is

van bijzonder belang dat men zich realiseert als het gaat om de relatie tussen de huisartsgeneeskundige en de specialistische setting. Voor huisartsen is de specialistische geneeskunde immers nog steeds de belangrijkste bron van kennis over symptomatologie. In dit artikel zullen wij deze materie toelichten aan de hand van een rekenvoorbeeld, waarbij een onderdeel van een vorig artikel nader wordt uitgewerkt.<sup>9</sup> Het gekozen voorbeeld heeft betrekking op het verband tussen moeheidsklachten en anemie. Er zijn echter aanwijzingen dat hetzelfde kan gelden voor andere klachten in relatie tot bepaalde ziekten, zoals het verband tussen dyspepsie en galstenen, en tussen hoofdpijn en hypertensie.<sup>10,11</sup>

## Anemie in de algemene bevolking

Uit diverse onderzoeken in de algemene bevolking is gebleken, dat er niet of nauwelijks een verband bestaat tussen matige anemie (c.q. een hemoglobineconcentratie onder een bepaalde waarde) en de „klassieke” klachten van anemie, zoals moeheid, duizeligheid, pijn op de borst en dyspnoe.<sup>12,13</sup> In de leerboeken staat moeheid echter beschreven als een van de belangrijkste symptomen van anemie, en ook in het onderwijs wordt hieraan aandacht besteed. Het ligt dan ook voor de hand dat huisartsen bij mensen met de klacht moeheid relatief vaak het Hb bepalen.

Uit de gegevens van het Diagnostisch Centrum Maastricht blijkt bijvoorbeeld dat bij de patiënten met de klacht moeheid die op aanvraag van de huisarts worden onderzocht, in 84 procent van de gevallen een Hb wordt bepaald. Bij de meeste andere klachten gebeurt dit veel minder vaak. Bij onderzoek in verband met een algemene screening of *check up* wordt slechts in 20 procent van de gevallen een Hb aangevraagd.<sup>14</sup>

In ons voorbeeld gaan we uit van de volgende – gefingeerde – basisgegevens:

- De kans op anemie is 5 procent:  $P(D) = 5$  procent.
- De klacht „moeheid” komt op een bepaald moment bij 10 procent van de algemene bevolking voor:  $P(S) = 10$  procent.
- Er is geen verband tussen moeheid en het voorkomen van anemie: de kans op anemie bij patiënten met moeheid is even groot als de kans op anemie bij mensen zonder moeheid, en de kans op moeheid bij anemie is even groot als de kans op moeheid bij afwezigheid van anemie. Dus:

$$P(S|D) = P(S|\bar{D}) = P(S), \\ \text{en} \\ P(D|S) = P(D|\bar{S}) = P(D).^{**}$$

In tabel 1 is de situatie in de algemene bevolking aangegeven, omgerekend naar 100.000 personen.

## Bij de huisarts

Het is aannemelijk dat mensen die moeheidsklachten hebben relatief vaak een beroep op de huisarts doen. Stel dat in een bepaalde periode 20 procent van de mensen met moeheidsklachten contact met de huisarts zoekt, en van de overige mensen 10 procent. Het Hb gehalte is niet bekend aan de patiënten, en is op zichzelf geen extra reden om contact met de huisarts zoeken. Daarom blijft het percentage personen met anemie in de groepen met respectievelijk zonder moeheidsklachten gelijk.

Uit tabel 2 blijkt nu dat in de populatie personen die een beroep doen op de huisarts (de *contactpopulatie*) moeheid vaker voorkomt dan in de algemene bevolking ( $P(S) = 18$  procent, in plaats van 10 procent). Er is echter nog steeds geen verband tussen moeheid en anemie:  $P(D|S) = P(D|\bar{S}) = 5$  procent. Nu komt er medisch handelen in het spel.

Doordat de patiënten aan de dokter laten merken dat zij weleens hebben gehoord dat moeheid door bloedarmoede kan worden veroorzaakt, en de dok-

\* Capaciteitsgroep Epidemiologie en Gezondheidszorgonderzoek van de Rijksuniversiteit Limburg.

\*\*  $P(X)$ : de kans op X. D = Disease (hier: anemie); S = Symptom (hier: moeheid).  $P(X|Y)$ : de kans op X, gegeven Y.

### Berkson's fallacy

In het kort komt deze vertekening op het volgende neer. Men vraagt zich af of factor F een (mede)oorzaak is van ziekte A. Om dit te onderzoeken gaat men na, of factor F bij in het ziekenhuis opgenomen patiënten met ziekte A vaker wordt aangetroffen dan bij een controle-groep, bestaande uit opgenomen patiënten met ziekte B. Het kan nu zijn dat de kans dat patiënten met ziekte A in het ziekenhuis opgenomen worden, positief wordt beïnvloed door de aanwezigheid van factor F. Stel dat een dergelijke invloed van factor F veel minder sterk of afwezig is ten aanzien van de kans op opname van patiënten met ziekte B. In dat geval zal men factor F, alleen al op grond van dit verschil in opnamekan-sen, vaker kunnen aantreffen bij patiënten met ziekte A dan bij patiënten met ziekte B. Een in feite niet aanwezig oorzakelijk verband tussen factor F en ziekte A kan hierdoor worden gesuggereerd of sterker lijken dan het in feite is. Op dezelfde wijze kan ook een beschermende werking van een bepaalde factor ten onrechte worden gesuggereerd. Als voorbeeld van dit laatste noemt *Feinstein* het onderzoek van *Pearl* uit 1929, die concludeerde dat tuberculose een antagonistische werking heeft ten opzichte van kanker. Hierop volgde een serie pogingen om met behulp van tuberculine kanker te bestrijden.<sup>7 8</sup>

ter dit bovendien heeft geleerd, zal bij patiënten met moeheid vaak een Hb worden bepaald - laten we zeggen: in 80 procent van de gevallen. Bij spreekuur-bezoekers zonder moeheid geschiedt dit veel minder vaak: in ons voorbeeld bij 8 procent van de gevallen. In tabel 3 zien we wat er nu gebeurt. Hoewel er geen verband is tussen moeheid en anemie, wordt een veel groter deel van de patiënten met anemie in de moeheidsgroep „ontdekt” dan in de groep zonder moeheid. Dit heeft tot gevolg dat van de anemiepatiënten die

„ontdekt” worden, ruim tweederde deel moeheidsklachten heeft.

Voor al de huisarts bij het opbouwen van zijn ervaringskennis alleen of hoofdzakelijk afgaat op de symptoma-tologie van ontdekte gevallen, zal een sterk verband tussen afwijking en symptoom worden gesuggereerd. Het verband is echter louter het produkt van de eigen indicatiestelling. Om begrijpelij-ke redenen zullen negatieve bevindingen in de dagelijkse praktijk minder aandacht krijgen. Praktijkervaring kan dan bedriegelijk zijn.

### Referentiegroepen

Huisartsen die deze valkuil willen ver-mijden, zullen wellicht besluiten tot het verrichten van een vergelijkend onder-zoek. Men moet dan wel een juiste controle- of referentiegroep kiezen.

Stel om te beginnen dat gekozen zou worden voor een vergelijking van de ontdekte anemiepatiënten met een steekproef uit de overige patiënten die het spreekuur bezoeken. Bij beide groe-pen wordt nagegaan hoe vaak er moeheidsklachten zijn gepresenteerd. De indruk zou nu bevestigd worden dat er een sterk verband bestaat tussen moeheid en anemie: in vergelijking met de anderen, hebben ontdekte anemiepa-tiënten in dit rekenvoorbeeld bijna viermaal zo vaak moeheidsklachten, na-melijk 69 tegenover 18 procent. Dit blijft vrijwel ongewijzigd het geval, als men de referentiegroep als geheel een Hb-onderzoek laat ondergaan om daar-mee andere anemiegevallen van de ver-gelijking uit te sluiten (controlegroepen a en b in tabel 4).

Een nog sterkere vertekening is het resultaat als de huisarts besluit om als referentiegroep een *random* steekproef uit de praktijkpopulatie (c.q. de alge-mene bevolking, inclusief de niet-spreekuurbezoekers) te kiezen, al dan niet met uitsluiting van patiënten met anemie: bij de ontdekte anemiegevallen komt nu bijna zevenmaal zo vaak moeheid voor: 69 tegenover 10 procent (controlegroepen c en d in tabel 4).

**Tabel 1. Rekenvoorbeeld over het vóórkomen van anemie en moeheid in de algemene bevolking.**

	Anemie (D)	Geen Anemie (D-)	Totaal
Moe (S)	500	9.500	10.000
Niet moe (S-)	4.500	85.500	90.000
Totaal	5.000	95.000	100.000

$$P(S) = P(S|D) = P(S|D-) = 10\%.$$

$$P(D) = P(D|S) = P(D|S-) = 5\%.$$

**Tabel 2. Het vóórkomen van anemie en moeheid in de populatie personen die de huisarts consulteert.**

	Anemie (D)	Geen Anemie (D-)	Totaal
Moe (S)	100	1.900	2.000
Niet moe (S-)	450	8.550	9.000
Totaal	550	10.450	11.000

$$P(S) = P(S|D) = P(S|D-) = 18\%.$$

$$P(D) = P(D|S) = P(D|S-) = 5\%.$$

**Tabel 3. De ontdekking van anemie door de huisarts, in relatie tot de klacht moeheid. Bij patiënten met moeheid wordt in 80% van de gevallen een Hb bepaald, bij patiënten zonder moeheid 8%.**

	Moe (S) (n = 2000)		Niet moe (S-) (n = 9000)	
	Anemie (D)	Geen anemie (D-)	Anemie (D)	Geen anemie (D-)
Bepaald	80	1520	36	684
Niet bepaald	20	380	414	7866
Totaal	100	1900	450	8550

Uitgaande van ontdekte anemiegevallen vindt men:

$$P(S|D) = \frac{80}{80 + 36} = 69\%.$$

### Toelichting

In tabel 1 wordt uitgegaan van de situatie dat er geen verband is tussen anemie en moeheid. Het voorbeeld heeft betrekking op een aselecte steekproef uit de algemene bevolking.

Van de mensen met moeheid bezoekt 20 procent de huisarts, van de mensen zonder moeheid 10 procent. Dit is weergegeven in tabel 2.

Vervolgens wordt bij patiënten met moeheid in 80 procent van de gevallen een Hb bepaald, en bij patiënten zonder moeheid in 8 procent van de gevallen (tabel 3).

Uitgaande van de oorspronkelijke groep patiënten die de huisarts bezoekt én een Hb-onderzoek krijgt, zouden in dit geval twee juiste vergelijkingen zijn te maken (tabel 5).

- Het vergelijken van de percentages anemiepatiënten onder de onderzochte patiënten mét respectievelijk zonder moeheidsklachten (beide 5 procent =  $80/1600 = 36/720$ ).

- Het vergelijken van de percentages personen met moeheidsklachten bij onderzochte patiënten mét en zonder anemie (beide 69 procent =  $80/116 = 1520/2204$ ).

Het is duidelijk dat het bij deze vergelijkingen om dezelfde subgroepen patiënten gaat, die echter in een andere richting met elkaar worden vergeleken.

Andere benaderingen voor het onderzoeken van het verband tussen de klacht moeheid en het Hb vanuit de huisartspraktijk zijn:

- Het trekken van een *random* steekproef van spreekuurbezoekers. Bij hen wordt geïnformeerd naar eventuele moeheidsklachten en tegelijkertijd wordt het Hb bepaald. Op deze wijze wordt vermeden dat de indicatiestelling door de arts vertekening veroorzaakt. Een beperking van deze methode is dat niet alleen die moeheidsklachten worden gemeld die tot doktersbezoek hebben geleid (*iatrotrope* klachten<sup>1)</sup>), maar ook allerlei andere wellicht minder ernstige moeheidsklachten. Niet uitgesloten is dat iatrotrope klachten in het algemeen anders van aard zijn dan niet-

iatrotrope (bijvoorbeeld via een enquête opgespoorde) klachten, ook wat betreft hun eventuele verband met het Hb. Voor de huisarts zijn met name de iatrotrope klachten van belang als startpunt van het handelen.

- Het systematisch bij alle patiënten die met moeheid het spreekuur bezoeken, het Hb bepalen en dit ook doen bij een referentiegroep van willekeurige spreekuurbezoekers zonder moeheid, en met klachten of problemen die geacht worden niet met het Hb samen te hangen.

Dit laatste is van belang, omdat het goed mogelijk is dat in de groep personen zonder moeheid, maar mét een kenmerkende indicatie voor een Hb-bepaling, meer anemie voorkomt dan bij „niet-

**Tabel 4.** *Vergelijkend onderzoek vanuit de huisartspraktijk naar het verband tussen anemie en het voorkomen van moeheid, op basis van de gegevens in de tabellen 1 en 2, bij de keuze van verkeerde referentiegroepen.*

	Percentage moe	Berekeningswijze
Door de huisarts ontdekte „anemiegevallen”	69,0%	$(= \frac{80}{116})$
a Alle spreekuurbezoekers exclusief de ontdekte anemiegevallen	17,6%	$(= \frac{1520+380+20}{11.000-116})$
b Alle spreekuurbezoekers exclusief alle anemiegevallen	18,2%	$(= \frac{1520+380}{11.000-(116+434)})$
c De praktijkbevolking exclusief de ontdekte anemiegevallen	9,9%	$(= \frac{10.000-80}{100.000-116})$
d De praktijkbevolking alle anemiegevallen	10,0%	$(= \frac{10.000-500}{100.000-5000})$

**Tabel 5.** *De correcte vergelijking met betrekking tot het verband tussen moeheid en anemie, uitgaande van de ontdekte „anemiegevallen”: de referentiegroep wordt gevormd door de onderzochte personen met een normaal Hb-gehalte.*

	Anemie (D)	Geen anemie (D-)	Totaal
Moe (S)	80	1520	1600
Niet moe (S-)	36	684	720
Totaal	116	2204	2320

$P(S|D) = P(S|D-) = 69\%$ .  $P(D|S) = P(D|S-) = 5\%$ .

**Tabel 6.** *Berekening van het percentage patiënten dat moeheidsklachten heeft in diverse subgroepen van de verwezen patiënten.*

Hb bepaald door huisarts	Moe		Niet moe	
	anemie	geen anemie	anemie	geen anemie
ja	80 p	1520 q	36 p	684 q
neen	20 q	380 q	414 q	7866 q

p - verwijsperscentage in geval van door de huisarts ontdekte anemie.  
q - verwijsperscentage bij de overige patiënten.

#### Percentage moeheid in diverse subgroepen

a Door huisarts ontdekte anemiepatiënten  $\frac{80p}{80p + 36p} = 69,0\%$

b Overige verwezen patiënten  $\frac{1920q}{1920q + 8964q} = 17,6\%$

c Overige verwezen patiënten met uitzondering van de na verwijzing nieuw ontdekte anemiepatiënten  $\frac{1900q}{1900q + 8550q} = 18,2\%$

d Door huisarts ontdekte anemiepatiënten én de na verwijzing nieuw ontdekte anemiepatiënten tezamen  $\frac{80p + 20q}{116p + 434q} = \text{afhankelijk van verwijfsfracties p en q}$

geïndiceerden". Een eventueel verschil zou dan kunnen worden verdoezeld. Om ruis te voorkomen, zou men de moeheidsgroep nog kunnen reduceren tot die groep patiënten bij wie een Hb-bepaling in de praktijk van belang (c.q. geïndiceerd) wordt geacht. Voor de hand liggende verklaringen voor de moeheid, anders dan anemie (bijvoorbeeld overbelasting) dienen dan eerst uitgesloten te worden.

## Verwijzing I

De huisarts zal vaak proberen de patiënten bij wie anemie is gevonden, te behandelen, bijvoorbeeld met ijzertherapie. Uit onderzoek is echter bekend dat dit in een aantal gevallen wel het Hb doet stijgen, maar het verloop van de moeheidsklachten niet beïnvloed.<sup>12</sup> Stel dat de huisarts een bepaald deel van de patiënten met "anemie" verwijst naar de specialist, bijvoorbeeld de internist. Wij bezien allereerst de - minst reële - situatie, waarin de kans op verwijzing alleen wordt beïnvloed door de gevonden anemie, en niet door het al dan niet moe zijn.

In deze situatie geldt dat de verhouding moe/niet moe onder de ontdekte anemiegevallen door verwijzing niet wordt beïnvloed. Dus van de door de huisarts ontdekte en verwezen patiënten met een onverklaarbare anemie, zal nog steeds ruim tweederde deel moeheidsklachten hebben en het overige eenderde deel niet (a in tabel 6). Dit betekent dat de specialist, evenals de huisarts, afgaande op zijn ervaring dat anemie vaak met hardnekkige moeheid gepaard gaat, gesteund zal worden in de opvatting dat anemie vaak met hardnekkige moeheid gepaard gaat.

Ook de specialist kan nu door middel van vergelijkend onderzoek proberen na te gaan, of zijn indruk dat anemiepatiënten relatief vaak moe zijn, juist is. Voor de specialist is het daarbij veel moeilijker dan voor de huisarts om een juiste referentiegroep te vinden. Men zou weer een steekproef kunnen nemen uit de lokale algemene bevolking. Omdat hiervan slechts 10 procent moeheidsklachten heeft, zal - ten onrechte - een sterk verband worden gesuggereerd. We beschreven deze valkuil al ten aanzien van de huisarts. Een iets minder grote vertekening zou optreden, indien als referentiegroep gekozen wordt voor (een steekproef uit) de overige verwezen patiënten, bij wie de huisarts geen anemie heeft ontdekt en die om andere redenen zijn verwezen. Voor

het gemak zou men deze groep kunnen laten representeren door één bepaalde categorie patiënten, die niet verwezen zijn in verband met een gevonden laag Hb, noch in verband met moeheid, bijvoorbeeld verwezen patiënten die preoperatief worden onderzocht in verband met een pogoperatie. In een dergelijke referentiegroep zal moeheid in ongeveer 18 procent van de gevallen voorkomen. Daarbij maakt het opnieuw niet veel uit of de na verwijzing bij routine-onderzoek alsnog ontdekte anemiegevallen wel of niet worden uitgesloten van de referentiegroep (b en c in tabel 6).

Dit alles betekent dat er in geval van vergelijking met de verwezen anemiepatiënten nog steeds een duidelijk verband tussen anemie en moeheid wordt gevonden.

Juiste benaderingen zouden zijn:

- Teruggaan naar de bron, dat wil zeggen opsporen van alle patiënten die bij de huisartsen een Hb-onderzoek kregen, inclusief de niet verwezen patiënten. Men zou nu, indien men het Hb aan het al dan niet gepresenteerd hebben van de klacht moeheid koppelt, tabel 5 weer kunnen opstellen.

- Terugprojecteren van de bron. Hier toe moet men weten of kunnen schatten:

- de fractie van de ontdekte anemiepatiënten, die door de huisarts is verwezen (p);

- de fractie van de overige patiënten die is verwezen (q).

Naast de reeds ontdekte anemiepatiënten worden nu door de specialisten ook de overige verwezen patiënten (of een representatieve steekproef uit die groep) op anemie gescreend, en van allen wordt genoteerd of de klacht moeheid is gepresenteerd. De pas na verwijzing ontdekte anemiegevallen worden bij de anemiepatiënten gerekend. Met behulp van de verwijsfracties en de absolute aantallen verwezen patiënten per categorie kan nu een correcte vergelijking worden ge(re)construeerd tussen alle anemiepatiënten en de referentiegroep (tabel 7).

- Het vinden van een (poli)klinische referentiegroep in de naar de specialist verwezen patiëntenpopulatie die naar aanleiding van de klacht moeheid door de huisarts precies even sterk geselecteerd is, als de door de huisarts ontdekte en verwezen anemiepatiënten.

Het is duidelijk dat deze benaderingen in de praktijk zeer moeilijk en dikwijls niet uitvoerbaar zullen zijn. Het achter-

af terugvinden van de bron van Hb-onderzoeken verricht door de huisarts met indicaties, klachten en uitslagen, is in de regel niet mogelijk. Hetzelfde geldt voor de schatting van de relevante verwijsfracties. Het achteraf kiezen van een qua moeheid vergelijkbaar geselecteerde referentiegroep zal in het algemeen neerkomen op oncontroleerbare veronderstellingen op dit punt.

Kunnen specialisten vanuit hun situatie dan helemaal geen voor de huisarts relevante waarnemingen doen met betrekking tot het verband tussen symptomen en aandoeningen?

Dat kan, en daartoe moeten zij uitgaan van de verwezen patiënten bij wie de huisarts voor de verwijzing geen Hb heeft bepaald, en een vergelijking maken tussen de volgende twee percentages:

- het percentage van de patiënten met anemie dat de klacht moeheid presenteert;

- het overeenkomstige percentage bij de patiënten met een normaal Hb.

Uit tabel 8 blijkt dat nu inderdaad geen verband tussen moeheid en een laag Hb wordt gevonden. Dit komt doordat de groepen die nu met elkaar vergeleken worden, via verwijzing in gelijke mate ten aanzien van de klacht moeheid zijn geselecteerd (niet beïnvloed door kennis over de Hb-waarde). Voor het kunnen maken van deze vergelijking is weer nodig dat ook de patiënten die verwezen zijn om andere redenen dan een laag Hb (of een steekproef van hen), na verwijzing een Hb-onderzoek krijgen.

Het is tot nog toe niet gebruikelijk om het verband tussen symptomen en diagnoses te onderzoeken bij patiënten die juist niet voor de betreffende diagnose zijn verwezen. Blijkens onze analyse is dit ten onrechte. Op dit gebied zouden specialisten en ziekenhuizen weleens een belangrijke bijdrage kunnen leveren aan de ontwikkeling van huisartsgeneeskundige kennis inzake het verband tussen klachten en ziekten. De contactpopulatie van de specialist is nu eenmaal beter bereikbaar voor wetenschappelijk onderzoek. De meet- en onderzoeksmethoden zijn gemakkelijk te standaardiseren, er is minder interobservervariatie te verwachten, en er zijn organisatorische voordelen. Zelf passen wij deze benadering op dit moment toe in een onderzoek naar het verband tussen dyspepsie en galstenen.

Een probleem bij deze methode is, dat vaak niet bekend zal zijn of de huisarts vóór de verwijzing wel of niet diagnostiek heeft verricht ten aanzien van de



**Tabel 7.** Berekening van het voor verwijzing gecorrigeerd percentage patiënten met moeheid in de groep anemiepatiënten.

- Stel dat na verwijzing de volgende verdeling van anemiegevallen in relatie tot moeheid wordt gevonden:

	Moe	Niet moe
Door de huisarts ontdekte en verwezen anemie	A	B
Na verwijzing ontdekte anemie	C	D

- Dan geldt dat de verdeling van de subgroepen *spreekuurbezoekers van de huisarts*, waaruit deze patiënten voortkomen, als volgt is:

	Moe	Niet moe
Door de huisarts ontdekte anemie	A/p	B/p
Door de huisarts niet ontdekte anemie	C/q	D/q

- Dan is het percentage moeheidspatiënten bij *alle* patiënten met anemie in de oorspronkelijke spreekuurpopulatie van de huisarts (= voor verwijzing gecorrigeerd percentage moeheid bij anemiepatiënten):

$$\frac{A/p + C/q}{A/p + B/p + C/q + D/q}$$

- In het voorbeeld is dit gelijk aan:

$$\frac{\frac{80p}{p} + \frac{20q}{q}}{\frac{116p}{p} + \frac{434q}{q}} + 18.2\%$$

- Dit is weer gelijk aan het percentage moeheidspatiënten in de categorie overige verwezen patiënten, met uitzondering van na verwijzing nieuw-ontdekte anemiepatiënten (zie tabel 6c):

p = verwijfspercentage in geval van door de huisarts ontdekte anemie.

q = verwijfspercentage bij de overige patiënten.

betreffende aandoening (in ons voorbeeld: het bepalen van het Hb). Wanneer het vóór verwijzing reeds verrichte onderzoek met uitslagen, ook als deze negatief zijn, systematisch in verwijfsbrieven zou worden vermeld, is dit echter op te lossen. Eventueel kan, in het kader van een onderzoek, additionele informatie betreffende de reeds door de huisarts verrichte diagnostiek worden ingewonnen.

Als de specialist volgens deze benadering te werk zou gaan zal het vooral om die ziektegevallen gaan, waarbij de huisarts geen indicatie heeft gesteld tot diagnostiek ten aanzien van de ziekte. Dit zullen in het algemeen minder evidente, en minder ontwikkelde gevallen zijn. Het zijn juist deze gevallen waarover het minst bekend is, en kennis hierover is vooral van belang voor de diagnostiek van de huisarts.

## Verwijzing II

De tweede verwijfssituatie – waarin de kans op verwijzing ook wordt beïnvloed door het hebben van hardnekkige moeheidsklachten – is het meest reëel, omdat hardnekkige moeheid een extra reden kan zijn om iemand met een laag Hb te verwijzen.

In deze situatie is te verwachten dat er onder anemiepatiënten naar verhouding nog vaker moeheid zal worden gevonden. Omdat nu de ontdekking van anemie én de klacht moeheid een rol spelen bij de selectie ten gevolge van de verwijzing, zijn de in de vorige paragraaf beschreven onderzoeksbenaderingen nog veel moeilijker toe te passen, zowel door hun complexiteit als door de benodigde gegevens. Zo zal voor elk van de vier mogelijke combinatie van klachtenpatroon (moe/niet moe) en Hb-

**Tabel 8.** Vergelijking van het percentage patiënten dat moeheidsklachten heeft binnen de groep overige verwezen patiënten (met en zonder anemie) bij wie vóór verwijzing door de huisarts geen Hb is bepaald (gegevens ontleend aan tabel 6).

	Anemie	Geen anemie
Moe	20q	380q
Niet moe	414q	7866q

Percentage moe bij deze patiënten:

$$\text{met anemie: } \frac{20q}{434q} = 4,6\%$$

$$\text{zonder anemie: } \frac{380q}{8246q} = 4,6\%$$

waarde (wel/niet anemie) een eigen verwijskans gelden.

Ook nu is het echter zo dat binnen de categorie patiënten, die niet in verband met anemie verwezen zijn en bij wie voor de verwijzing geen Hb is bepaald, het verband zonder vertekening te bestuderen is. De binnen deze categorie alsnog te identificeren anemiepatiënten zijn namelijk ten aanzien van moeheid weer even sterk geselecteerd als de te identificeren patiënten zonder anemie.

In werkelijkheid zijn er nog meer problemen dan we bespraken. Van de diverse subcategorieën van de groep „overige” verwezen patiënten zullen in de regel geen gegevens bekend zijn over hun specifieke verwijsfracties. Meestal zal men ook te weinig weten over een eventueel verband van de betreffende diagnoses en hun verwijsfracties enerzijds met het Hb en/of de klacht moeheid anderzijds. Ook de invloed van ernst en de aard van de klacht of het klinisch beeld op de verwijzing is onvoldoende bekend. Alleen prospectief onderzoek in samenwerking met huisartsen, waarbij uitgegaan wordt van de contactpopulatie van de huisarts, kan de problemen wat betreft de opgetreden selectie naar de specialist toe definitief oplossen. Zoals gezegd, kan onderzoek bij patiënten die verwezen zijn om andere redenen dan de te bestuderen diagnose, een aanvaardbaar alternatief bieden.

## Beschouwing

In het voorgaande is aangetoond dat een bepaald ingesleten of aangeleerd diagnostisch gedrag ten aanzien van het verband tussen symptomen/klachten en diagnostische bevindingen een *self-fulfilling prophecy*-effect te weeg kan brengen. Het betreffende verband kan dan,



ook als het in werkelijkheid niet bestaat. worden aangetroffen, met name als men zijn empirische ervaring vooral baseert op positieve bevindingen. In de huisartspraktijk zijn dergelijke indrukken door een goed vergelijkend onderzoek nog wel te corrigeren. Uitgaande van een populatie verwezen patiënten, is dit veel moeilijker, en vaak onmogelijk.

Op allerlei niveaus kunnen er selectiemechanismen werkzaam zijn die schijfverbanden opleveren. Achtereenvolgens kan men onderscheiden: de algemene bevolking, bezoek aan de huisarts, resultaten van diens handelen, verwijzing naar een (bepaald) specialisme, interspecialistische verwijzingen (bijvoorbeeld van kleinere ziekenhuizen naar grotere, voorzien van superspecialismen), ziekenhuisopname, bevestiging van diagnoses door middel van pathologisch anatomisch onderzoek of sectie. Bij elk van de tussenstappen van het ene naar het andere niveau wordt geselecteerd op grond van hetgeen men heeft geleerd.

Schijnverbanden kunnen ook ontstaan ten gevolge van het publiceren in bekende tijdschriften van *case-reports* betreffende bijzondere klinische waarnemingen, zoals een nog niet eerder beschreven combinatie van symptomen. Zo'n *case-report* kan aanleiding geven tot het verzamelen van alle nieuwe of in het verleden niet „herkende“ gevallen met dit symptomencomplex. Op grond van zo'n patiëntenserie kan dan een nieuw syndroom in het leven worden geroepen. Vaak is de noemer van de epidemiologische breuk waarvan deze serie gevallen de teller is, niet meer bekend of krijgt deze geen aandacht, evenmin als de wijze van selectie van deze gevallen. Toevallige combinaties kunnen dan de status van syndroom krijgen, en worden door alerte waarnemers (en wie wil dat niet zijn) daarna weer regelmatig gezien. Het lijkt interessant om van een aantal beschreven syndromen zonder duidelijke verklaring (zoals de Trias van Saint: galstenen-diverticulosis-hiatus hernia) na te gaan, of een dergelijk mechanisme aan de orde kan zijn. Vertekende verbanden kunnen weer verpakt worden in medisch onderwijs, of op klinische bijeenkomsten als belangrijke diagnostische aangrijppingspunten voor huisartsen naar voren worden gebracht.

En zo wordt de cirkel gesloten. Men creëert wat men doceert, men doceert wat men creëert.

Het is al met al erg belangrijk dat men telkens de bevindingen op een bepaald

zorgniveau (echelon) op de juiste wijze kan relateren aan de relevante patiëntenbron. Dit impliceert voor de huisartsgeneeskunde:

- het verrichten van onderzoek, juist vanaf het niveau van de huisartspraktijk;
- op grond daarvan eventueel het bijstellen van de onderwezen kennis over bijvoorbeeld symptomatologie;
- het verrichten van onderzoek naar de samenhang van klachten en ziekten bij patiënten die verwezen zijn om andere redenen dan de betreffende ziekten.

We maken nog enkele algemene opmerkingen.

Men kan zeggen: het gaat om het stellen van een diagnose en niet om de precieze weg waarlangs dat gebeurt. Als daarbij gebruik wordt gemaakt van vertekende verbanden, is dat op zichzelf niet erg. Deze redenering klopt niet. Als men het stellen van de diagnose inderdaad relevant vindt, dient men zich niet te baseren op weinig of niet-informatieve diagnostische gegevens. Als deze gebruikt worden om te selecteren voor nader diagnostisch onderzoek, zullen in elk geval veel diagnoses gemist worden. Men had dan net zo goed een dubbelsteen kunnen opwerpen, of men had iedereen aan het betreffende onderzoek moeten onderwerpen om het missen van veel diagnoses te vermijden. Men kan ook zeggen: de verbanden die zich in de loop der jaren in het medisch kennissysteem hebben uitgekristalliseerd, zijn naar voren gekomen vanuit een in oorsprong „blanco“ situatie. Dat wil zeggen een situatie waarin artsen niet *gebiased* waren door onderwijs op basis van een vertekende waarneming. De verbanden zullen in het algemeen dus wel kloppen. Deze bewering zou opgaan als er in de tijd niets veranderde ten aanzien van het verband tussen symptomen en diagnoses. Dit ligt echter niet voor de hand. Ziektebeelden (hun ernst, de wijze waarop ze zich manifesteren en de diagnostische criteria) zijn wel aan veranderingen in een tijd onderhevig, evenals hun verhouding tot andere, deels nieuwe aandoeeningen en ziektebeelden. Bovendien is symptomatologie, behalve historisch, ook cultureel bepaald: de klachten vermagering en moeheid hebben in Nederland anno 1985 een heel andere betekenis dan in de Derde Wereld, of in Nederland anno 1885.

Er is ongetwijfeld een zekere „traagheid“ in de ontwikkeling van de kennis over symptomatologie. Symptomatologie uit vervlogen tijden kan ons in de

han houden en ingesleten routines wortelen haast per definitie het sterkst in een al gepasseerd tijdperk.

Dat leidt tot nog een conclusie: het is van belang, van tijd tot tijd bepaalde waarheden op het gebied van de symptomatologie nog eens te onderzoeken, ook als dit in het verleden reeds correct is gebeurd. Dit onderzoek zou vooral gericht moeten worden op aspecten die gevoelig zijn voor veranderingen in de tijd. De hier beschreven selectiefenomenen gelden bovendien ook in omgekeerde zin: als men bestaande verbanden niet kent, zal men ze ook niet gauw op het spoor komen. De ontdekking van minder sterke verbanden is zeer afhankelijk van de beschikbaarheid van oplettende waarnemers. Voor de huisartspraktijk kunnen dat alleen huisartsen zijn.

<sup>1</sup> Berkson J. Limitations to the application of fourfold table analysis to hospital data. *Biometrics Bull* 1946; 2: 47-54.

<sup>2</sup> Rümke Cl. Gevaren voor onjuiste conclusies uit gegevens in ziekenhuisarchieven (Berkson's fallacy). *Ned Tijdschr Geneesk* 1970; 114: 766-72.

<sup>3</sup> Boyd AV. Testing for association of diseases. *J Chron Dis* 1979; 32: 667-72.

<sup>4</sup> Sturmans F. Epidemiologie, theorie, methoden en toepassing. Nijmegen: 1982. Dekker & Van de Vegt.

<sup>5</sup> Peritz E. Berkson's bias revisited. *J Chron Dis* 1984; 37: 909-16.

<sup>6</sup> Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic research, principles and quantitative methods. Belmont (Ca): Wadsworth, 1982.

<sup>7</sup> Feinstein AR. Clinical biostatistics. Saint Louis: Mosby, 1977.

<sup>8</sup> Pearl R. Cancer and tuberculosis. *Am J Hyg* 1929; 9: 97-159.

<sup>9</sup> Knottnerus JA. Interpretatie van diagnostische gegevens. *Huisarts en Wetenschap* 1983; 26: 363-8.

<sup>10</sup> Knipschild PG. Dyspepsie en galstenen. *Practitioner* (Ned uitgave) 1984; 1: 1126-9.

<sup>11</sup> Knipschild PG, Vissers A, Knottnerus A. Hoofdpijn en hoge bloeddruk. Uitkomsten van een bevolkingsonderzoek in Amsterdam. *Huisarts en Wetenschap* 1984; 27: 441-5.

<sup>12</sup> Elwood PC, Waters WF, Greene WJW, Sweetnam P, Wood MM. Symptoms and circulating haemoglobin level. *J Chron Dis* 1969; 21: 615-28.

<sup>13</sup> Wood MM, Elwood PC. Symptoms of iron deficiency anemia, a community survey. *Br J Prevent Soc Med* 1966; 20: 117-21.

<sup>14</sup> Knottnerus A, Knipschild PG. Hemoglobine en „anaemische“ klachten zoals moeheid. *Tijdschr Soc Gezondheidszorg* 1983; 61: 788-92.

<sup>15</sup> Feinstein AR. Clinical judgement. New York: Kruger, 1967.

## DEEL II

### ENKELE OBSERVATIONELE ONDERZOEKINGEN NAAR HET VERBAND TUSSEN KLACHTEN EN DIAGNOSTISCHE BEVINDINGEN.

# Hoofdpijn en hoge bloeddruk

Uitkomsten van een bevolkingsonderzoek in Amsterdam

PAUL KNIPSCHILD, AD VISSERS EN ANDRÉ KNOTTNERUS\*

Het is nog steeds niet duidelijk, of hoofdpijn een voorspellende betekenis heeft voor de aanwezigheid van hoge bloeddruk. In dit artikel wordt, met behulp van de uitkomsten van een bevolkingsonderzoek onder 3695 Amsterdammers van 41-43 jaar, onderzocht of een dergelijk verband bestaat. De frequentie en ernst van de hoofdpijn werden afgemeten aan het gebruik van middelen tegen hoofdpijn. Bij de 199 deelnemers die deze middelen frequent gebruikten, was de bloeddruk gemiddeld 128/81 mm Hg, tegen 130/81 mm Hg bij de overige deelnemers. Een diastolische bloeddruk van 105 mm Hg en meer werd iets vaker gevonden in de hoofdpijngroep dan in de controlegroep.

## Inleiding

Mensen met hoge bloeddruk lopen een belangrijk risico om bepaalde ziekten te krijgen, die voornamelijk worden veroorzaakt door de overbelasting en daarvoor verslechterende functie van het arteriële vaatstelsel. Aangenomen mag worden dat het risico wordt verkleind door tijdige opsporing en verlaging van de hoge bloeddruk. Daarom worden voortdurend pogingen ondernomen om mensen met hoge bloeddruk te ontdekken. Een probleem daarbij is dat deze mensen zich niet snel bij een arts zullen melden, omdat hoge bloeddruk meestal geen klachten geeft.

Of toch wel? Veel artsen hebben geleerd dat hoge bloeddruk nogal eens gepaard gaat met bepaalde, meestal vage klachten. Vooral hoofdpijn wordt vaak genoemd als een begeleidend verschijnsel en om die reden wordt hoofdpijn wel beschouwd als een indicatie om de bloeddruk te meten.<sup>1</sup> Uit de literatuur (*kader* op pagina 444) is evenwel niet duidelijk op te maken, of er inderdaad een verband bestaat tussen hoofdpijn en hoge bloeddruk en hoe sterk dat eventuele verband is. Daarmee blijft tevens de vraag onbeantwoord, in hoeverre hoofdpijn een diagnostische betekenis heeft voor het opsporen van hoge bloeddruk. Onlangs ontstond de gelegenheid om deze kwestie nogmaals te bestuderen.

## Methode

De gegevens werden verzameld in het kader van een bevolkingsonderzoek ter opsporing van risicofactoren voor hart- en vaatziekten in Amsterdam, het zogeheten Consultatie Bureau Project. De afdeling Bevolkingsregister van de gemeente verstrekke een lijst van alle personen, woonachtig in Amsterdam en geboren in 1937 (bewoners van de Westelijke tuinsteden 1936). In de periode november 1977 tot september 1980 ontvingen 2422 mannen en 2748 vrouwen, allen in de leeftijd van 41-43 jaar, een uitnodiging om zich te laten screenen.

Van de mannen nam 11 procent en van de vrouwen 4 procent niet aan het onderzoek deel, omdat men onlangs elders al gescreeend was op deze risicofactoren (vooral via het COMI-project in de bedrijven). Bovendien deed 3 procent van de mannen en 2 procent van de vrouwen niet mee, omdat men al onder behandeling was wegens een hartziekte. Tenslotte nam 19 procent van de mannen en 18 procent van de vrouwen niet aan het onderzoek deel zonder opgave van redenen. Daarmee bestond de onderzoekspopulatie uit 1647 mannen (68 procent) en 2048 vrouwen (75 procent);

de non-participatie bedroeg 29 procent. Aan alle deelnemers werd een aantal mondelinge vragen gesteld, waarna enkele metingen werden verricht. Voor de onderhavige vraagstelling waren vooral twee vragen en één meting van belang: • *Hoofdpijn*. In de vragenlijst was de volgende vraag opgenomen: „Gebruikt u wel eens middelen tegen hoofdpijn (aspirine of dergelijke)? Zo ja, doet u dit zelden, iedere maand, iedere week, iedere dag, in periodes?” Helaas werd niet directer naar hoofdpijnklaarten gevraagd. Voor onze analyse is een driedeling gemaakt: personen die zelden of nooit hoofdpijnmiddelen gebruikten, personen die iedere maand (of in periodes) hoofdpijnmiddelen gebruikten en personen die iedere week (of zelfs iedere dag) hoofdpijnmiddelen gebruikten.

• *Hoge bloeddruk-anamnese*. In de vragenlijst was de volgende vraag opgenomen: „Als bij u ooit een verhoogde bloeddruk gevonden? Gebruikt u hiervoor geneesmiddelen? Zo ja, welke?” Aldus was het mogelijk om de relatie tussen hoofdpijn en hoge bloeddruk apart te bestuderen voor personen met een bekende c.q. onbekende hoge bloeddruk.

• *Bloeddrukmeting*. Bij alle deelnemers werd de bloeddruk op gestandaardiseerde wijze gemeten met een gewone, regelmatig geijkte kwik-manometer. Dit gebeurde terwijl de deelnemers gezeten waren, met ontbloot bovenlijf. Indien meer metingen waren verricht, is voor de analyse steeds de eerste meting genomen. Als systolische bloeddruk werd de waarde genoteerd, waarbij voor het eerst ritmisch doorgaande tonen werden gehoord. Als diastolische bloeddruk werd de waarde aangehouden, waarbij net geen tonen meer hoorbaar waren, behalve als de tonen hoorbaar bleven tot 0 mm Hg of dicht daarbij; in dat geval werd het moment van duidelijk zachter worden van de tonen genoteerd.

## Resultaten

Alvorens in te gaan op de vraag, of hoofdpijn een voorspellende betekenis heeft voor de hoogte van de bloeddruk, wordt eerst gekeken naar de mogelijkheid van iatrogene hoofdpijn.

Uit tabel 1 blijkt dat 604 van de 3695 deelnemers meldden dat bij hen ooit een verhoogde bloeddruk was gevonden. In deze groep gebruikte 5 procent (29 personen) iedere week middelen tegen hoofdpijn en 80 procent (486 personen) zelden of nooit. Bij 327 van de

\* De auteurs zijn verbonden aan de capaciteitsgroep Epidemiologie van de Rijksuniversiteit Limburg.

604 deelnemers kan men zich afvragen, of ze wel in de groep met een positieve anamnese thuishoren, omdat ze bij meting geen hoge bloeddruk (meer) bleken te hebben en ook geen bloeddrukverhogende medicatie gebruikten. In de groep van 277 personen met een „te-recht positieve” anamnese gebruikte eveneens 5 procent (15 personen) iedere week middelen tegen hoofdpijn en 79 procent (218 personen) zelden of nooit. Indien het bekend zijn met hoge bloeddruk hoofdpijn kan veroorzaken, mag in de groep met een negatieve anamnese een lager gebruik van middelen tegen hoofdpijn worden verwacht. Dat blijkt echter niet het geval. In deze groep van 3091 deelnemers gebruikte ook 5 procent (170 personen) iedere week middelen tegen hoofdpijn en 79 procent (2450 personen) zelden of nooit. Ook in de groep met wel aanwezige, maar nog niet ontdekte hoge bloeddruk (272 personen) werden nagenoeg dezelfde percentages gevonden: 6 procent (15 personen) gebruikte iedere week middelen tegen hoofdpijn en 82 procent (224 personen) zelden of nooit.

Uit tabel 1 komen dus geen aanwijzingen voor een verband tussen een positieve anamnese ten aanzien van hoge bloeddruk en het gebruik van middelen tegen hoofdpijn. Indien dit gebruik indicatief is voor de frequentie en ernst van hoofdpijn, geeft het onderzoek dus geen aanwijzingen voor het bestaan van iatrogene hoofdpijn.

Voor het beantwoorden van de vraag, of mensen met hoofdpijn vaker hoge bloeddruk hebben, hoeft men de deelnemers met een positieve anamnese ten aanzien van hoge bloeddruk dus niet buiten beschouwing te laten. Cijfers over deze vraag zijn uit dezelfde tabel 1 te halen, door de tabel niet in horizonta-

**Table 1 – Table 1. Het gebruik van middelen tegen hoofdpijn bij 3695 personen – Use of anticephalalgics in 3695 persons.**

History of hypertension	Medication	Hypertension at examination*	Use of anticephalalgics			Total
			seldom/never	monthly	weekly	
+	+	+	55	8	7	70
+	+	–	47	17	2	66
+	–	+	116	19	6	141
+	–	–	268	45	14	327
–	–	+	224	33	15	272
–	–	–	2226	438	155	2819
Total			2936	560	199	3695

**Table 2 – Table 2. Bloeddruk en gebruik van middelen tegen hoofdpijn bij 3695 personen. Blood pressure and use of anticephalalgics in 3695 persons.**

	Use of anticephalalgics		
	seldom/never (n = 2936)	monthly (n = 560)	weekly (n = 199)
<b>Systolic blood pressure (mm Hg)</b>			
Mean	129.9	128.3	127.8
Standard deviation	17.0	16.4	16.5
< 100	44 (1%)	6 (1%)	0 (0%)
100-139	2235 (76%)	458 (82%)	152 (76%)
140-159	473 (16%)	70 (13%)	37 (19%)
160-179	160 (5%)	20 (4%)	9 (5%)
180-199	20 (1%)	4 (1%)	1 (1%)
≥ 200	4 (0%)	2 (0%)	0 (0%)
<b>Diastolic blood pressure (mm Hg)</b>			
Mean	80.9	80.9	81.0
Standard deviation	11.4	10.7	12.2
< 60	36 (1%)	9 (2%)	2 (1%)
60-89	2345 (80%)	460 (82%)	160 (80%)
90-94	226 (8%)	41 (7%)	12 (6%)
95-99	136 (5%)	14 (3%)	8 (4%)
100-104	112 (4%)	24 (4%)	7 (4%)
105-109	29 (1%)	4 (1%)	4 (2%)
110-114	31 (1%)	5 (1%)	3 (2%)
115-119	7 (0%)	2 (0%)	1 (1%)
≥ 120	14 (0%)	1 (0%)	2 (1%)

\* Blood pressure systolic ≥ 160 mm Hg and/or diastolic ≥ 95 mm Hg.

## Synopsis

**Knipschild P, Visser A, Knotterus A. Headache and hypertension. Results of a population study in Amsterdam. Huisarts en Wetenschap 1984; 27: 441-5.**

**Introduction.** The authors discuss the possibility of an association between headache and hypertension.

**Methods.** Data were analyzed of a population study among 3695 inhabitants of Amsterdam, all being 41 to 43 years old. The

frequency and severity of headache was measured through the use of anticephalalgics. Questions were also asked about a positive history of hypertension. The blood pressure was measured in all participants. **Results.** Table 1 gives no indication for an association of the use of anticephalalgics with a positive history of hypertension, nor with measured hypertension. Table 2, however, reveals a relationship between a high diastolic blood pressure (≥ 105 mm Hg) and the use of anticephalalgics. This pattern is confirmed in tables 3 and 4.

**Discussion.** This association is statistically significant, but so far it seems to be of little clinical importance. Further investigation is required: to what extent is unexplained headache indicative of hypertension, and what is the relationship between migraine and hypertension?

**Keywords:** Headache. Hypertension. Population study.

**Address of authors:** Capaciteitsgroep Epidemiologie, Rijksuniversiteit Limburg, Postbus 616, 6200 MD Maastricht.

le, maar in verticale richting te lezen. Aldus ontstaat het volgende beeld: Van alle deelnemers meldde 199 personen dat ze iedere week middelen te-

gen hoofdpijn gebruikten en in deze groep werd bij 14 procent (28 personen) een hoge bloeddruk gemeten. Nage- noeg dezelfde percentages werden ge-

**Tabel 3 - Table 3.** Het voorkomen van een diastolische bloeddruk  $\geq 105$  mm Hg in relatie tot het gebruik van middelen tegen hoofdpijn - The prevalence of diastolic blood pressure  $\geq 105$  mm Hg in relation to the use of anticephalalgics.

	Use of anticephalalgics		
	seldom/ never	monthly	weekly
<b>Men</b>			
Positive history	20 ( 9%)	3 (12%)	2 (17%)
Negative history	21 ( 2%)	1 ( 1%)	2 ( 4%)
Total	41 ( 3%)	4 ( 2%)	4 ( 6%)
<b>Women</b>			
Positive history	28 (11%)	4 ( 6%)	2 (12%)
Negative history	12 ( 1%)	4 ( 1%)	4 ( 4%)
Total	40 ( 3%)	8 ( 2%)	6 ( 5%)
<b>Men and women</b>			
Positive history	48 (10%)	7 ( 8%)	4 (14%)
Negative history	33 ( 1%)	5 ( 1%)	6 ( 4%) <sup>a</sup>
Total	81 ( 3%)	12 ( 2%)	10 ( 5%) <sup>b</sup>

<sup>a</sup> Weekly versus seldom/never + monthly;  $p = 0.02$ , Fisher's exact test with modification according to Miettinen.

<sup>b</sup> Weekly versus seldom/never + monthly;  $p = 0.04$ , Fisher's exact test with modification according to Miettinen.

**Tabel 4 - Table 4.** Bloeddruk en gebruik van middelen tegen hoofdpijn bij 3091 personen met een negatieve anamnese ten aanzien van hoge bloeddruk. Blood pressure and use of anticephalalgics in 3091 persons with a negative history of hypertension.

	Use of anticephalalgics		
	seldom/ never (n = 2450)	monthly (n = 471)	weekly (n = 170)
<b>Systolic blood pressure (mm Hg)</b>			
Mean	127.6	126.4	126.0
Standard deviation	15.9	15.2	15.1
< 100	43 ( 2%)	5 ( 1%)	0 ( 0%)
100-139	1980 (81%)	404 (86%)	138 (81%)
140-159	331 (14%)	48 (10%)	27 (16%)
160-179	85 ( 3%)	10 ( 2%)	4 ( 2%)
188-199	9 ( 0%)	4 ( 1%)	1 ( 1%)
$\geq 200$	2 ( 0%)	0 ( 0%)	0 ( 0%)
<b>Diastolic blood pressure (mm Hg)</b>			
Mean	79.3	79.5	79.5
Standard deviation	10.5	9.9	11.3
< 60	36 ( 1%)	9 ( 2%)	2 ( 1%)
60-89	2075 (85%)	405 (86%)	145 (85%)
90-94	156 ( 6%)	30 ( 6%)	10 ( 6%)
95-99	90 ( 4%)	10 ( 2%)	3 ( 2%)
100-104	60 ( 2%)	12 ( 3%)	4 ( 2%)
105-109	9 ( 0%)	2 ( 0%)	4 ( 2%)
110-114	15 ( 1%)	2 ( 0%)	0 ( 0%)
115-119	4 ( 0%)	1 ( 0%)	0 ( 0%)
$\geq 120$	5 ( 0%)	0 ( 0%)	2 ( 1%)

vonden voor de groep deelnemers die iedere maand middelen tegen hoofdpijn gebruikte (60/560 = 11 procent) en voor de groep die deze middelen zelden of nooit gebruikte (395/2936 = 13 procent).

In tabel 2 zijn de gemiddelde bloeddrukwaarden genoteerd. Ook deze cijfers geven geen aanwijzingen dat hoofdpijn, afgemeten aan het gebruik van middelen tegen hoofdpijn, een voorspellende betekenis voor de bloeddruk heeft. Verder blijken zeer hoge systolische bloeddrukken niet vaker te worden aangetroffen in de groep die iedere week middelen tegen hoofdpijn gebruikt. Voor de diastolische bloeddruk ligt dat anders. Hoe hoger het afkappunt gekozen wordt, des te meer discrimineert het gebruik van middelen tegen hoofdpijn. Tabel 3 heeft alleen betrekking op de deelnemers met een diastolische bloeddruk  $\geq 105$  mm Hg. Van alle 199 personen die wekelijks middelen tegen hoofdpijn gebruikten, had 5 procent (10 personen) zo'n hoge bloeddruk. Het overeenkomstige percentage onder de overige deelnemers bedraagt bijna 3 procent. Het verschil is statistisch significant. Hoofdpijn, althans het gebruik van middelen tegen hoofdpijn, lijkt dus enige voorspellende betekenis te hebben voor de aanwezigheid van een zeer hoge diastolische bloeddruk.

Voor de volledigheid zijn in de tabellen 1, 3 en 4 ook cijfers gegeven voor de groep deelnemers bij wie nooit eerder een verhoogde bloeddruk is gevonden. Hoge bloeddruk, gedefinieerd als een bloeddruk systolisch  $\geq 160$  en/of diastolisch  $\geq 95$  mm Hg, lijkt niet vaker voor te komen in de groep die iedere week middelen tegen hoofdpijn gebruikt (tabel 1). Ook de gemiddelde bloeddruk is in deze groep niet hoger (tabel 4). Wel blijkt nu ook weer dat zeer hoge diastolische bloeddrukken iets vaker in deze groep worden gemeten (tabel 3).

## Beschouwing

De gegevens waarvan hier gebruik is gemaakt, zijn destijds verzameld met het oog op een geheel andere vraagstelling: bovendien is onze vraagstelling pas geformuleerd, toen de gegevens al binnen waren. We mogen dus aannemen dat het antwoord op de vraag naar het gebruik van middelen tegen hoofdpijn geen invloed heeft gehad op de bloeddrukmeting.

Uit onze secundaire analyse kwamen geen aanwijzingen dat het bekend raken met hoge bloeddruk hoofdpijn kan veroorzaken. Het onderzoek was echter

vooral gericht op de vraag, of hoge bloeddruk zelf hoofdpijn kan veroorzaken. In de literatuur wordt hoofdpijn vooral vermeld als een symptoom van maligne hypertensie, met een diastolische bloeddruk van minstens 120 mm Hg. Uit dit onderzoek komen aanwijzingen dat ook in de groep mensen met een diastolische bloeddruk van 105-119 mm Hg iets vaker middelen tegen

hoofdpijn worden gebruikt. Uit tabel 2 blijkt dat 9 procent van de 86 personen met een dergelijke bloeddruk iedere week middelen tegen hoofdpijn gebruikt, terwijl dat in de grote groep van 3592 personen met een lagere bloeddruk bij 5 procent het geval is. Van de 44 personen bij wie voor het eerst een diastolische bloeddruk van minstens 105 mm Hg werd gemeten,

gebruikten zes personen iedere week middelen tegen hoofdpijn (sensitiviteit = 14 procent). In de grote groep van 3047 personen met een lagere diastolische bloeddruk was dit bij 164 personen het geval (100%-specificiteit = 5 procent). Het onderscheidend vermogen, gedefinieerd als sensitiviteit/100%-specificiteit, bedraagt daarmee 2,5. Omdat diastolische bloeddrukken van minstens

## De literatuur

De eerste ons bekende publikatie over het verband tussen hoge bloeddruk en hoofdpijn dateert van zeventig jaar geleden. *Janeway* beschrijft een serie van 458 patiënten met een systolische bloeddruk boven 160 mm Hg en meldt: „Cerebral symptoms were noted early by many patients. Headache was the most frequent. (...) So commonly have patients described to me a particular kind of headache that I have almost come to look on it as a typical nephritic or hypertension symptom. This headache is one which appears on awakening, or wakes the patient during the early morning hours, has its greatest intensity before arising, and passes away either immediately after breakfast or during the course of the morning”.

De beschreven hypertensie-hoofdpijn werd aangetroffen bij 13 procent van alle patiënten. Opvallend daarbij is dat dergelijke hoofdpijn even vaak leek voor te komen in de groep met een systolische bloeddruk van 165-220 mm Hg als in de groep met een hogere bloeddruk. Het is overigens nog maar de vraag, of bij mensen zonder hoge bloeddruk dergelijke hoofdpijn veel minder voorkomt. Daarover geeft de publikatie van *Janeway* geen uitsluitsel. Wil men tot een uitspraak komen over de relatie tussen hoofdpijn en hoge bloeddruk, dan is er een controlegroep nodig: uitgaande van de vraag, of bij mensen met hoofdpijn vaak hoge bloeddruk voorkomt, dient men tevens de frequentie van hoge bloeddruk te schatten bij een groep mensen zonder hoofdpijn. In de meeste publikaties heeft men daarbij gemakshalve de volgorde omgekeerd. Indien hoofdpijn een voorspellende betekenis heeft voor de aanwezigheid van hoge bloeddruk, zal bij hoge bloeddruk-patiënten vaak hoofdpijn worden aangetroffen. Men selecteert een groep mensen met hoge bloeddruk en een controlegroep zonder hoge bloeddruk en vraagt naar het optreden van hoofdpijn.

Hoewel de omkering op zichzelf niet tot onjuiste conclusies hoeft te leiden, ontstaat nu wel het gevaar van selectie-bias: niet alle hoge bloeddruk wordt ontdekt en als de ontdekking mede veroorzaakt wordt door de aanwezigheid van hoofdpijn, kunnen schijnverbanden ontstaan. Deze selectie van hoge bloeddruk-patiënten kan vooral optreden, als patiënten met hoge bloeddruk bij aanwezigheid van hoofdpijn eerder naar de specialist worden verwezen, en als men vervolgens een groep hoge bloeddruk-patiënten bij de

specialist op de klacht hoofdpijn vergelijkt met een controlegroep zonder hoge bloeddruk, bijvoorbeeld in de algemene bevolking. Het is dan niet verwonderlijk dat men bij patiënten met nieuw ontdekte hoge bloeddruk in een hypertensiekliniek vaker hoofdpijn aantreft.

Bij een dergelijke onderzoekspopet ontstaat nog een extra mogelijkheid van bias, als men zich niet beperkt tot patiënten met een nieuw ontdekte hoge bloeddruk, maar vooral patiënten met een al langer bekende hoge bloeddruk in het onderzoek betreft (de prevalentie gevallen). In de literatuur wordt gesuggereerd dat bekendheid met het hebben van hoge bloeddruk invloed heeft op het ziektegevoel. Klachten als hoofdpijn zouden kunnen ontstaan, nadat men is ingelicht over zijn hoge bloeddruk en de bijbehorende risico's.<sup>4</sup> Kiest men nu in het onderzoek voor prevalentie gevallen van hoge bloeddruk<sup>5</sup>, dan kan het gevonden verband tussen hoofdpijn en hoge bloeddruk nog eerder tot een verkeerde interpretatie leiden.

De vraag of hoofdpijn een reden is om aan hoge bloeddruk te denken, kan goed onderzocht worden in de huisartspraktijk. Tot dusver zijn hierover twee publikaties verschenen. Walker onderzocht zijn praktijkpopulatie op het verband tussen migraine en bloeddruk. Het gaat dus niet over hoofdpijn in het algemeen of ochtendhoofdpijn, maar over aanvallen van hoofdpijn die in ieder geval af en toe unilateraal optreden niet daarbij eventueel visuele verschijnselen en misselijkheid. Uit zijn onderzoek bleek onder meer dat bij migrainepatiënten de bloeddruk systolisch en diastolisch zo'n 10 mm Hg hoger was dan bij patiënten zonder migraine.<sup>6</sup> Deze uitkomst is interessant in het licht van de observatie van *Janeway*, dat „a surprisingly large number of these patients (met hoge bloeddruk en hoofdpijn) has been subject to migraine throughout life”.<sup>2</sup>

In een ander onderzoek in de huisartspraktijk ging het wel om hoofdpijn in het algemeen. Vergelijken werden de frequentie en ernst van hoofdpijn bij een groep mensen bij wie door de huisarts voor het eerst hoge bloeddruk werd ontdekt, en bij een controlegroep waarbij de bloeddruk niet te hoog bleek te zijn. Tussen beide groepen werd geen verschil aangetroffen.<sup>7</sup>

Helaas zijn in dit artikel enkele dingen onduidelijk. Hoe werd de controlegroep geselecteerd? En bij welke bloeddruk werd ge-

sproken van hoge bloeddruk? Daarmee is het nog steeds mogelijk dat mensen die wegens (bepaalde) hoofdpijn de huisarts consulten, vaker hoge bloeddruk hebben.

Tenslotte is er een aantal publikaties, waarin het verband tussen hoofdpijn en hoge bloeddruk bestudeerd is op het niveau van de algemene bevolking. Het aantrekkelijke van deze aanpak is dat allerlei selectiemechanismen in mindere mate optreden. Dergelijk onderzoek is vooral geschikt om te kijken, of hoofdpijn inderdaad een symptoom is van hoge bloeddruk. Een nadeel van onderzoek in de algemene bevolking is dat grote aantallen mensen onderzocht moeten worden om uitspraken te kunnen doen over het verband tussen hoofdpijn en zeer hoge bloeddruk. Bovendien is er het dilemma of de mensen met reeds eerder gediagnosticeerde hoge bloeddruk wel of niet in de vergelijking moeten worden betrokken. Doet men dit wel, dan is vertekening denkbaar door eventuele iatrogene hoofdpijn. Laat men daarentegen alle eerder ontdekte gevallen van hoge bloeddruk uit de vergelijking weg, dan is het zelfs mogelijk dat men een negatief (schijn-) verband vindt. Stel het theoretische geval dat alle mensen met hoofdpijn al op hoge bloeddruk zijn gescreend en vervolgens alle ontdekte gevallen (met de combinatie van beide) niet meer mee mogen doen. In dat geval wordt in de geselecteerde subpopulatie natuurlijk geen hoge bloeddruk meer gevonden in de groep mensen met hoofdpijn.

Kennelijk was een aantal onderzoekers niet zo bang voor deze selectie-bias, want men ging uit van een populatie waarbij de mensen met (al langer) bekende hoge bloeddruk uit het onderzoek werden gelaten. In de groep mensen met hoofdpijn werd vervolgens niet vaker hoge bloeddruk gevonden.<sup>8-10</sup> In twee andere publikaties houdt men zich meer op de vlakte over de samenstelling van de onderzoeksgroep. In één daarvan wordt een duidelijk verband gevonden tussen hoofdpijn en hoge bloeddruk, terwijl bij hypertensiepatiënten ook meer migraine voorkomt.<sup>11</sup> Uit de andere publikatie komen wederom geen aanwijzingen voor een verband tussen hoofdpijn, migraine en hoge bloeddruk.<sup>12</sup> Tenslotte wordt hier een recente publikatie vermeld, waaruit blijkt dat de frequentie van ernstige hoofdpijn nauwelijks hoger is bij een groep met een onbehandelde, zeer hoge diastolische bloeddruk, vergeleken met groepen met iets lagere bloeddrukken.<sup>13</sup>

105 mm Hg echter maar weinig worden aangetroffen (in dit onderzoek bij ruim 1 procent van alle deelnemers met een negatieve anamnese), is de extra informatie met betrekking tot het symptoom hoofdpijn gering. Voor het besluit om wel of niet de bloeddruk te meten ter opsporing van een zeer hoge diastolische bloeddruk, lijkt hoofdpijn dus vooralsnog nauwelijks een leidraad.

Hoe moet het gevonden verband tussen hoofdpijn en een hoge diastolische bloeddruk worden geïnterpreteerd? Vooralsnog wordt ervan uitgegaan dat dit verband geen toevalstreffer is. Dat hoofdpijn (als uiting van „stress“) een oorzaak zou zijn van een hoge diastolische bloeddruk, is niet erg aannemelijk. Men zou dan zeker ook een relatie met de systolische bloeddruk verwachten en bovendien is het nog maar zeer de vraag, of chronische „stress“ de bloeddruk doet stijgen.<sup>14</sup> Er is evenmin reden om te veronderstellen dat bepaalde middelen tegen hoofdpijn (phenacetine) voor het verband verantwoordelijk zijn. In het beschreven onderzoek is niet gevraagd naar type hoofdpijnmiddel. Analgetica-nefropathieën zijn echter in de algemene bevolking zo zeldzaam dat ook aldus geen afdoende verklaring gegeven kan worden. Veel meer moet daarom gedacht worden aan een mechanisme waarbij de hoge diastolische bloeddruk (via een invloed op de liquor-druk of op de cerebrale arteriolen?) hoofdpijn kan veroorzaken. Interessant is dan bovendien de vraag, of bij patiënten met „hypertensie-hoofdpijn“ deze klacht minder wordt, als de bloeddruk verlaagd wordt.

Hier wordt vooral ingegaan op een andere kwestie. In hoeverre is hoofdpijn een reden om hoge bloeddruk te vermoeden en dus de bloeddruk te meten? Uit de gepresenteerde resultaten blijkt dat er wel een verband is tussen hoofdpijn en een (zeer) hoge diastolische bloeddruk. Tegelijkertijd blijkt echter dat het wel of niet hebben van hoofdpijn nauwelijks discrimineert tussen hoge en lagere bloeddrukken. Dat kan te maken hebben met de weinig nauwkeurige wijze waarop het hebben van hoofdpijn is vastgesteld. Het is denkbaar dat nogal wat mensen in de groep die zelden of nooit middelen tegen hoofdpijn zegt te gebruiken, wel frequent lijden onder

ernstige hoofdpijn. In dat geval zullen de verschillen in relatie tot hoge bloeddruk in werkelijkheid iets groter zijn dan uit de gepresenteerde cijfers blijkt. Veel meer „ruis“ wordt mogelijk veroorzaakt door een andere onvolkomenheid van dit onderzoek (en de meeste andere onderzoeken). Het is bekend dat veel vormen van hoofdpijn, gepaard gaand met frequent gebruik van middelen hiertegen, niets te maken hebben met hoge bloeddruk. Hoofdpijn hangt vaak samen met bepaalde gedragspatronen, zoals te veel alcohol drinken en te weinig nachtrust. Daarnaast bestaat er een relatie met bepaalde somatische aandoeningen, zoals sinusitis en hersenschudding. In de gepresenteerde analyse kon helaas de hoofdpijn door algemeen geaccepteerde oorzaken niet buiten beschouwing worden gelaten. Het is denkbaar dat onverklaarde hoofdpijn wel duidelijker discrimineert tussen hoge en lagere bloeddrukken.

Nieuw onderzoek is zinvol met als vraagstelling, of onverklaarde hoofdpijn indicatief is voor de aanwezigheid van hoge bloeddruk. Wellicht dient daarbij ook een specificatie gemaakt te worden naar de lokalisatie en het tijdstip van optreden van de hoofdpijn. Er wordt immers in de literatuur gesuggereerd dat „hypertensie-hoofdpijn“ vooral in het achterhoofd zit en in de vroege ochtend optreedt. Ook is hernieuwde aandacht gewenst voor het verband tussen migraine en hoge bloeddruk.

Om selectie-bias te vermijden kan dit onderzoek beter niet op specialistisch niveau plaatsvinden. Voor de hand ligt om het onderzoek opnieuw in de algemene bevolking te verrichten, maar het is zeker ook aantrekkelijk in de huisartspraktijk. Globaal zou dergelijk onderzoek er als volgt kunnen uitzien: De huisarts meet bij ieder contact met een patiënt met hoofdpijn de bloeddruk op gestandaardiseerde wijze, bij voorkeur met een geblindeerde, random-zero meter. De totale groep onderzochte patiënten wordt opgesplitst in patiënten met onverklaarde hoofdpijn (onderzoeksgroep 1) en met migraine (onderzoeksgroep 2) en de bloeddruk van deze patiënten wordt vergeleken met die van andere patiënten met een geaccepteerde oorzaak van de hoofdpijn (controlegroep 1). Uit het aanbod

van patiënten zonder hoofdpijn wordt een extra controlegroep toegevoegd.

# Dankbetuiging

Onze dank gaat uit naar Hans Meijer en Herman Sallé van het Coronel Laboratorium van de Universiteit van Amsterdam, die een belangrijke rol hebben gespeeld bij het verwerken van de gegevens, en naar de medewerkers van het Consultatie Bureau Project Hartziekten in Amsterdam, die alle benodigde gegevens hebben verzameld.

- <sup>1</sup> Overdijk T. Huisarts en hypertensie. Naslagboek voor snelle oriëntatie. Utrecht: Nederlands Huisartsen Instituut, 1980.
- <sup>2</sup> Janeway TC. A clinical study of hypertensive cardiovascular disease. Arch Intern Med 1913; 12: 755-98.
- <sup>3</sup> Bulpitt CJ, Dollery CT, Carne S. Change in symptoms of hypertensive patients after referral to hospital clinic. Br Heart J 1976; 38: 121-8.
- <sup>4</sup> Stewart IMG. Headache and hypertension. Lancet 1953; i: 1261-6.
- <sup>5</sup> Badran RHA, Weir RJ, McGuinness JB. Hypertension and headache. Scott Med J 1970; 15: 48-51.
- <sup>6</sup> Walker CH. Migraine and its relationship to hypertension. Br Med J 1959; ii: 1430-3.
- <sup>7</sup> Robinson JO. Symptoms and the discovery of high blood pressure. J Psychosom Res 1969; 13: 157-61.
- <sup>8</sup> Weiss NS. Relationship of high blood pressure to headache, epistaxis, and selected other symptoms. The United States Health Examination Survey of Adults. N Engl J Med 1972; 287: 631-3.
- <sup>9</sup> Kottke TE, Tuomilehto J, Puska P, Salonen JT. The relationship of symptoms and blood pressure in a population sample. Int J Epidemiol 1979; 8: 355-9.
- <sup>10</sup> Reek J van, Diederiks J, Philipsen H, Zutphen W van, Seelen T. Subjective complaints and blood pressure. J Psychosom Res 1982; 26: 155-65.
- <sup>11</sup> Hofman O, Kolár M, Reisenauer R, Matousek V. Significance of the differences in the prevalence of the subjective complaints between normotensive and hypertensive subjects. Acta Univ Carol [Med] (Praha) 1973; 19: 601-16.
- <sup>12</sup> Waters WE. Headache and blood pressure in the community. Br Med J 1971; i: 142-3.
- <sup>13</sup> Kraus JF, Curb JD, Cutter G, Daugherty SA, Neill KC, Wassertheil-Smolter S. Baseline medical history characteristics of the hypertensive participants. Hypertension detection and follow-up program. Hypertension 1983; 5 (part II, no. 6): 51-91.
- <sup>14</sup> Anonymous. Genetics, environment, and hypertension [Editorial]. Lancet 1983; i: 681-2.



# Hemoglobine en 'anemische' klachten zoals moeheid

A. Knottnerus, dr. P. G. Knipschild<sup>1</sup>

Er zijn veel leemten in kennis over de zin en de betekenis van de zo vertrouwde hemoglobinebepaling, en het belang van het hemoglobinegehalte van het bloed voor de gezondheid. Een oriënterend onderzoek werd verricht naar de hemoglobinewaarden bij patiënten die op verzoek van de huisarts onderzocht waren: patiënten met de hoofdklacht moeheid ( $n=566$ ), en een controlegroep van personen die zonder speciale klacht een algemeen onderzoek ondergaan hadden ( $n=112$ ).

Er werd geen relevant verschil in hemoglobinewaarden gevonden. Alleen bij oudere mannen was het hemoglobine in de moeheidsgroep iets lager dan in de controlegroep. Ook literatuurgegevens suggereren weinig samenhang tussen hemoglobinegehalte en 'anemische' klachten zoals moeheid. De vraag of in de huisartspraktijk verschillende hemoglobinewaarden gevonden worden bij patiënten met en zonder deze klachten, is echter onvoldoende beantwoord. Omdat de huisarts het hemoglobinegehalte nogal eens bepaalt bij patiënten met moeheid, zal een vergelijkend onderzoek worden gedaan in huisartsenpraktijken naar de relatie tussen moeheid en hemoglobine.

T. Soc. Gezondheidsz. 61 (1983) 788-792

## Inleiding

In alle sectoren van de gezondheidszorg wordt vele malen daags de hemoglobinebepaling uitgevoerd. Met name binnen de diagnostiek van de huisarts neemt zij een belangrijke plaats in. Vaak wordt de bepaling gebruikt om in twijfelgevallen uit te maken of men al dan niet in somatische richting zal doorzoeken.

Het is echter opmerkelijk hoeveel leemten er bestaan in de kennis over de betekenis van het hemoglobinegehalte in het bloed (Hb) voor de gezondheid en de gezondheidszorg. Deze leemten betreffen zowel de bijdrage die de Hb-bepaling levert aan de ontdekking van verborgen pathologie en de daarbij te hanteren normaalwaarden, als ook de consequenties van een bepaald Hb-gehalte op lange termijn.

Voor dit laatste aspect is door de epidemiologie nog weinig bestudeerd. Waters en Elwood (1982) komen op grond van hun eigen onderzoek tot de conclusie dat weliswaar duidelijk verlaagde Hb-gehalten gepaard gaan aan een licht verhoogde sterftekans, maar dat een licht verlaagd Hb mogelijk samenhangt met een lagere mortaliteit aan cardiovasculaire ziekten. Een aanzienlijke stijging van het sterftecijfer is waargenomen bij personen met een Hb hoger dan 15 g%.<sup>2</sup> Deze tendenzen zijn nog onvoldoende onderzocht. Dat geldt ook voor de mogelijke verklaringen, zoals het frekwenter voorkomen van coronaire anastomosen bij anemiepatiënten (Zoll e.a. 1951) en de gevonden positieve correlatie tussen hemoglobinegehalte en cholesterolspiegel (Elwood e.a. 1970).

De normaalwaarden die aangehouden worden voor het Hb-gehalte, geven de arts weinig houvast. In het algemeen zijn zij vastgesteld op statistische en niet op klinische gronden, na onderzoek van klachtenvrije populaties. Zij zeggen dus niets over de gezondheidstoestand. Een probleem hierbij is dat het Hb een continue variabele is en een unimodale, bij benadering normale verdeling volgt. Het kiezen van een 'afkappunt' als caesuur voor normale respectievelijk afwijkende Hb-waarden is dan een arbitraire zaak. In dit verband is de benadering van Garby (1970) vermeldenswaard. Hij kon bij een steekproef van vrouwen uit de bevolking van Uppsala de verdeling van het Hb en de hematocriet<sup>3</sup> onderscheiden bij vrouwen met en zonder ijzergebrek. Dit geschiedde door middel van het bepalen van de respons op orale ijzertherapie. Met behulp van deze verdelingen werd het mogelijk om voor elke waarde van het Hb en de hematocriet de bijbehorende kans op ijzergebrek te berekenen.

Voor de praktiserend arts is het zinvol het probleem van de betekenis van de Hb-bepaling te benaderen vanuit de klachten van de patiënt. Daarin ligt immers in eerste instantie de indicatie tot het doen van de bepaling. Richtsnoer voor de arts is daarbij hetgeen deze geleerd heeft over de symptomen die kunnen optreden bij een laag hemoglobinegehalte en anemie. Dit artikel is vooral vanuit deze invalshoek geschreven. De centrale vraag daarbij is:

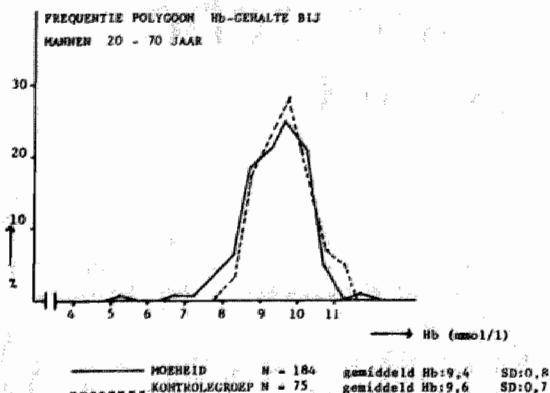
Hebben mensen met klachten, die per traditie heten samen te hangen met een laag Hb, ook inderdaad lagere Hb-waarden dan mensen die deze klachten niet hebben? Met andere woorden: geeft het doen van een Hb-bepaling bij mensen met deze klachten een differentiatie ten opzichte van hen die ze niet hebben?

<sup>2</sup> In de internationale literatuur wordt het Hb meestal nog vermeld in g%. Een g% komt overeen met 0,621 mmol/l.

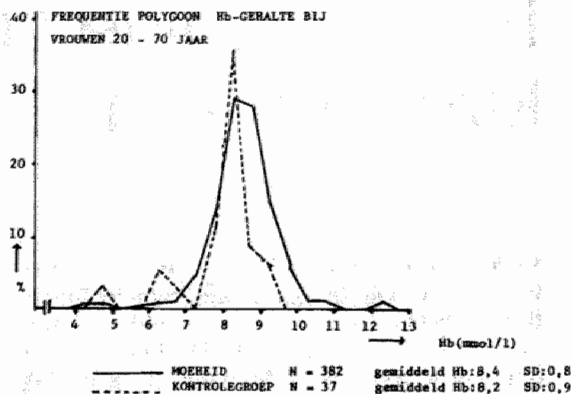
<sup>3</sup> De hematocriet is het totale volume van de rode bloedcellen per 100 ml bloed

<sup>1</sup> Arts resp. sociaal-geneeskundige, Capaciteitsgroep Epidemiologie van de Rijksuniversiteit Limburg





**Figuur 1.** Frekwentiepolygon van hemoglobineconcentraties in het bloed van mannen van 20-70 jaar, wegens de klacht moeheid, resp. in verband met een keuring of algemeen onderzoek (controlegroep) onderzocht op het diagnostisch centrum Maastricht



**Figuur 2.** Frekwentiepolygon van hemoglobineconcentraties in het bloed van vrouwen van 20-70 jaar, wegens de klacht moeheid, resp. in verband met een keuring of algemeen onderzoek (controlegroep) onderzocht op het diagnostisch centrum Maastricht

### Oriënterend onderzoek

Moeheid is volgens de gegevens van Lamberts (1982) na hoesten, keelontsteking en nervositeit de meest frekwente nieuwe klacht op het huisartsensprek uur. Deze klacht staat nog genoteerd vóór lage rug- en hoofdpijn. Aangezien moeheid vrij algemeen gezien wordt als een belangrijk symptoom van anemie, was onze verwachting dat deze klacht relatief vaak tot een Hb-bepaling zou leiden. Ter oriëntatie besloten wij tot een retrospectief onderzoekje in het databestand van het Diagnostisch Centrum Maastricht<sup>4</sup>.

### Methode

In het Diagnostisch Centrum wordt op aanvraag van de huisartsen uit de regio Maastricht en omstreken laboratoriumonderzoek verricht. Op het aanvraagformulier vermeldt de huisarts de klacht(en) en de reden van het onderzoek. In 1980 en 1982 werden voor evaluatiedoeleinden gedurende in totaal 6 maanden alle gegevens betreffende aanvragen en uitslagen in de computer ingevoerd. Daardoor was het mogelijk de Hb-uitslagen over deze maanden te vergelijken voor de volgende groepen:

- patiënten die voor de eerste keer voor bloedonderzoek waren verwezen wegens de hoofdklacht moeheid (moeheidsgroep),
- patiënten zonder speciale klacht (controlegroep). Zij waren naar het lab verwezen in verband met een keuring of een verzoek om een algemeen onderzoek. Binnen het beschikbare materiaal leek de kans op een van een klachten-vrije populatie afwijkende Hb-verdeling bij deze categorie het geringst.

Alle bepalingen op het laboratorium vonden plaats tussen 8.00 uur en 16.00 uur. Er is geen reden om te veronderstellen

dat de bepalingen voor beide groepen op systematisch verschillende tijdstippen verricht werden.

### Resultaten

De gegevens werden bewerkt voor de leeftijdsgroep van 20 tot en met 69 jaar. Onder de leeftijd van 20 jaar waren bij de mannen geen controlepersonen beschikbaar. Bovendien verandert het Hb gedurende de eerste 20 levensjaren sterk met de leeftijd, terwijl op zeer hoge leeftijd bij mannen een lichte daling van het hemoglobinegehalte plaatsvindt (Kelly en Munan 1977, Waters en Elwood 1982). Uiteindelijk telde de moeheidsgroep 382 vrouwen en 184 mannen, en de controlegroep 37 vrouwen en 75 mannen, die in de genoemde leeftijdscategorie vielen en bij wie een Hb bepaald was. Daarbij moet opgemerkt worden dat bij de totale groep mensen, die wegens moeheid verwezen waren, in een veel hoger percentage van de gevallen een Hb bepaald was (84%) dan bij de totale groep personen, die onderzocht was wegens een keuring of algemeen onderzoek (21%).

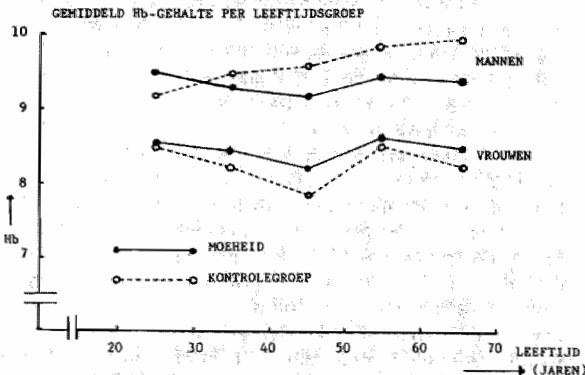
In de figuren 1 en 2 zijn de verdelingen gegeven voor beide groepen, voor mannen en vrouwen apart, met gemiddelden en standaardafwijking. Bij mannen is er sprake van een significant verschil tussen de gemiddelden ten nadele van de moeheidsgroep ( $p < .05$ ). Echter een verschil tussen de gemiddelden van 0,2 mmol/l, gevonden voor zowel mannen als vrouwen, kan men niet relevant noemen.

Een aanvullend beeld krijgen we, als we kijken naar tabel 1,

**Tabel 1.** Gemiddelden en standaarddeviaties van het hemoglobinegehalte per leeftijdscategorie voor de moeheidsgroep en de controlegroep, mannen en vrouwen van 20-70 jaar, in mmol/l

leeftijd	MANNEN				VROUWEN			
	moeheid n	Hb ± SA	controle n	Hb ± SA	moeheid n	Hb ± SA	controle n	Hb ± SA
20-29	56	9.5 ± 0.7	4	9.2 ± 0.7	109	8.5 ± 0.6	6	8.5 ± 0.4
30-39	43	9.3 ± 0.8	12	9.5 ± 0.5	90	8.4 ± 0.6	10	8.2 ± 0.5
40-49	35	9.2 ± 0.7	37	9.6 ± 0.7	83	8.2 ± 1.1	7	7.8 ± 1.9
50-59	30	9.4 ± 1.2	15	9.8 ± 0.8	54	8.6 ± 0.9	4	8.5 ± 2.9
60-69	20	9.3 ± 0.9	7	9.9 ± 0.9	46	8.5 ± 0.8	10	8.2 ± 0.8
<b>totaal</b>	<b>184</b>	<b>9.4 ± 0.8</b>	<b>75</b>	<b>9.6 ± 0.7</b>	<b>382</b>	<b>8.4 ± 0.8</b>	<b>37</b>	<b>8.2 ± 0.9</b>

<sup>4</sup> De data werden verkregen dankzij de medewerking van P. Kerkhof, arts, capaciteitsgroep huisartsgeneeskunde Rijksuniversiteit Limburg



Figuur 3. Gemiddeld hemoglobinegehalte per leeftijdscategorie naar geslacht, voor moeheidsgroep en controlegroep, in mmol/l.

gevisualiseerd in figuur 3. Het lijkt erop dat bij mannen bij toenemende leeftijd een steeds duidelijker verschil bestaat tussen beide groepen ten nadele van de moeheidsgroep, oplopend tot 0.6 mmol/l. Anders gezegd: bij oudere mannen lijkt de klacht moeheid samen te hangen met een lager Hb, konform de verwachting. Bij vrouwen daarentegen is het Hb in de controlegroep steeds iets lager dan in de moeheidsgroep.

Passen we de normaalwaarden van het Diagnostisch Centrum toe op de gegevens, dan ontstaat tabel 2. Bij mannen ligt een hoger percentage van de Hb's in de moeheidsgroep onder de norm. Dit beeld is niet terug te vinden bij de vrouwen.

Tabel 2. Aandeel van de Hb-waarden, dat onder de normaalwaarden van het diagnostisch centrum Maastricht ligt, in moeheidsgroep en controlegroep, voor mannen en vrouwen van 20-70 jaar.

	MANNEN (8.6 mmol/l)		VROUWEN (7.4 mmol/l)	
	fractie	percentage	fractie	percentage
moeheidsgroep	25/184	14 %	20/382	5 %
controlegroep	4/75	5 %	3/37	8 %

## Bespreking

In het beschreven onderzoek kwamen nauwelijks relevante verschillen naar voren in de verdeling van hemoglobine-waarden bij mensen die wegens moeheid een bloedonderzoek ondergingen, en degenen zonder deze klacht die onderzocht werden. Alleen bij oudere mannen komt het resul-

taat overeen met de verwachting.

Direkt moet hierbij worden aangetekend dat het materiaal beperkte waarde heeft. Het was niet voor dit doel verzameld, en voor de definiering van de klacht moeheid bestonden geen werkafspraken. Daarnaast is het mogelijk dat eventuele verschillen versluierd zijn, omdat bij de mensen die wegens moeheid naar het diagnostisch centrum verwezen zijn, veel vaker een Hb geprikt is, dan bij de mensen die gekeurd werden. Als de huisarts speciale - niet vermelde - redenen had om bij sommige keuringen wel een Hb aan te vragen en bij andere niet, kunnen bij de wel op Hb onderzochten naar verhouding immers meer lage waarden gevonden zijn. We hebben voorts geen informatie over de Hb-waarden van alle personen die het spreekuur van de huisarts bezochten wegens moeheid of voor een keuring, dat wil zeggen inclusief degenen die niet naar het diagnostisch centrum verwezen zijn. Tenslotte geldt de beperking, dat in enkele leeftijdsklassen van de controlegroep de aantallen erg klein waren.

Als we kijken naar de literatuur, blijkt dat er weinig onderzoek is gedaan naar de relatie tussen klachten en Hb-gehalte. We moeten ook ver teruggaan in de tijd.

Berry en Nash (1954) onderzochten 360 huisvrouwen, die deelnamen aan een routine-röntgenborstonderzoek en daarbij vrij bleken van tuberculose. Gevraagd werd of men zich fit en energiek voelde; vaak moe was, of kortademig. Tevens werd het Hb bepaald. Er werd geen verschil gevonden wat betreft het gemiddelde Hb tussen de vrouwen met en zonder klachten. Evenmin was er een significant verschil in de percentages vrouwen die een Hb hadden onder de 10,8 g%.

In 1966 gingen Wood en Elwood bij een steekproef uit de bevolking van Wales (165 mannen, 130 vrouwen) na, of er een correlatie bestond tussen klachten en hemoglobinegehalte. Bij vrouwen werd alleen voor (door henzelf opgemerkte) bleekheid een verband gevonden met een lager Hb. Bij mannen bleken kortademigheid en pijn op de borst samen te gaan met een hoger Hb. Voor moeheid, slapte, hartkloppingen, prikkelbaarheid en duizeligheid werd geen relatie met de hoogte van het Hb gevonden.

Vanuit hetzelfde gebied en door hetzelfde team werden enkele jaren later nieuwe resultaten gerapporteerd (Elwood e.a. 1969). Nu werden alleen vrouwen onderzocht, allen niet-zwanger en in de leeftijd van 20-64 jaar, 880 in getal. Er werd gevraagd naar de aanwezigheid dan wel de graad van ernst van de volgende klachten: moeheid, prikkelbaarheid, duizeligheid, hartkloppingen, kortademigheid en hoofdpijn, en los daarvan werd ook weer het Hb bepaald. Tussen klachten en Hb werd geen enkel verband gevonden.

In dezelfde publikatie wordt verslag gedaan van twee expe-

rimenten, waarbij het effect van ijzertherapie op genoemde klachten is onderzocht, zowel bij vrouwen met een Hb van 13,5 g% of hoger als bij vrouwen met een Hb tussen 8 en 12 g%. In beide categorieën werd geen positief effect van ijzertoeiening op de klachten gevonden ten opzichte van de placebogroep. Wel vond er in de categorie van 8-12 g% in de behandelde groep een sterkere Hb-stijging plaats dan in de placebogroep.

Een beperking van de genoemde studies is dat er te weinig personen gevonden werden met een zeer laag Hb (lager dan 8 g%) om een mogelijk verband tussen klachten en deze lage Hb-waarden apart te onderzoeken. Klinische observaties worden echter juist vaak gedaan bij mensen met sterk verlaagde Hb-gehalten. Elwood en Wood (1966) deden wel een onderzoek bij 20 ziekenhuispatiënten met Hb's variërend van 6 tot 10 g%, maar het gunstige effect dat ijzer en de gevonden Hb-stijging leken te hebben op hun klachtenpatroon is niet goed te beoordelen bij het ontbreken van een controlegroep.

Dawson e.a. (1969) stelden een onderzoek in bij ziekenhuispatiënten van 50 jaar en ouder. Zij vergeleken het optreden van klachten bij anemiepatiënten met een Hb van 8-10 g% (n=70) en anemiepatiënten met een Hb lager dan 8 g% (n=63). Alleen bleekheid werd frekwenter in de tweede groep gevonden. Voor de overige klachten, zoals moeheid, duizeligheid, hoofdpijn en hartkloppingen was er geen verschil. Wel kwamen er in absolute zin veel klachten voor. Zo klaagden in beide groepen ongeveer 90% van de patiënten over moeheid. Helaas wordt de geslachtsverdeling niet voor beide groepen apart verstrekt, en is op grond van de gegevens ook geen vergelijking tussen de groepen mogelijk wat betreft de onderliggende oorzaken van de anemie. Deze zouden namelijk, onafhankelijk van het Hb, genoemde klachten kunnen veroorzaken.

In de huisartspraktijk is nauwelijks vergelijkend onderzoek gedaan naar het verband tussen een laag Hb en klachten. Watt (1972) bepaalde het Hb bij 98 patiënten die zij op grond van klachten en klinisch beeld verdacht van anemie. Dezelfde bepaling werd verricht bij 99 willekeurig gekozen spreekuurbezoekers, die niet van anemie verdacht werden. De groepen waren vergelijkbaar wat betreft leeftijd en geslacht, doch bevatten nauwelijks mannen. In beide groepen werd een even hoog percentage Hb-waarden lager dan 12 g% gevonden. Echter, alleen in de controlegroep kwamen zwangeren voor, en wel 16 in getal. Een eventueel verschil zou daardoor kunnen zijn gemaskeerd.

De laatste tien jaar hebben diverse onderzoekers (Viteri en Torùn 1974, Edgerton e.a. 1979, Ohira e.a. 1981) een positief verband gevonden tussen Hb-gehalte en maximaal arbeidsvermogen. Daarbij is ook vastgesteld dat ijzertoe-

ning dit vermogen gunstig beïnvloedt bij mensen met ijzergebrek. Gegeven de biologische kennis, dat hemoglobine het transportmiddel is voor zuurstof naar de weefsels, is dit ook te verwachten. De genoemde onderzoekers vergaarden hun gegevens echter in de derde wereld. Het blijft de vraag in hoeverre het gevonden verband voor de medicus practikus in Nederland van belang is bij de diagnostiek. De voedingstoestand is hier veel beter en weinig mensen moeten gaan tot het maximum van hun fysieke kunnen. De bevindingen zijn mogelijk van meer belang voor de bedrijfs- en de sportgeneeskunde. In ieder geval geven zij een extra dimensie aan de discussie over normaalwaarden. Viteri en Torùn spreken over hematologische normaalwaarden en anemie als 'dynamische concepten waarbij rekening moet worden gehouden met de zuurstofbehoefte van het individu en diens mogelijkheden om daaraan tegemoet te komen'.

Resumerend kan gesteld worden, dat de vraag of bij mensen, die op het spreekuur komen met 'anemische' klachten zoals moeheid, de hemoglobinebepaling een differentiatie geeft ten opzichte van mensen zonder deze klachten, op grond van ons oriënterend onderzoekje en de literatuur niet afdoende kan worden beantwoord.

Op het niveau van de algemene populatie is er geen onderscheid gemaakt tussen spontaan geuite en door de onderzoekers voorgelegde klachten. Wat de huisartspraktijk betreft, is het onderzoek van Watt onbevredigend opgezet. In slechts één onderzoek, dat van Wood en Elwood (1966) in de algemene populatie, is gelet op eventuele verschillen tussen mannen en vrouwen wat betreft het verband tussen klachten en Hb-gehalte. Het is overigens mogelijk dat een dergelijk verschil, indien het zou bestaan, eerder te maken heeft met een ander patroon van oorzaken van een lager Hb dan met het lage Hb zelf. Zowel Kirschenfeld en Tew (1955) als Fry (1961) melden na onderzoek van grote series patiënten in de huisartspraktijk dat bij mannen in geval van anemie meestal sprake is van onderliggende ziekteprocessen, terwijl bij vrouwen vaak geen andere oorzaak wordt gevonden dan overmatig menstrueel bloedverlies, een deficiënt dieet, of zwangerschap.

### Nader onderzoek

Met name voor de huisarts is het van belang of de Hb-bepaling zinvol is. Veel mensen met 'anemische' klachten komen op diens spreekuur. Vaak spelen allerlei problemen daarbij een rol - op het gebied van relaties, gezin of werk - die mensen moe, duizelig of geprikkeld kunnen maken of hartkloppingen bezorgen. Als men dan een Hb bepaalt, zou heel wel bij toeval - en bij herhaling - een waarde kunnen worden gevonden die onder de gehanteerde norm valt.

Nader onderzoek dreigt dan, tot beenmergpunctie aan toe, vooral als een positieve respons op ijzertherapie uitblijft. Processen als somatische fixatie (Grol, 1981) krijgen aldus een kans. Aan de andere kant weet de huisarts dat ernstige pathologie weliswaar zelden voorkomt bij 'anemische' (vage) klachten, maar zich, indien aanwezig, juist vaak aandient met dergelijke klachten. De vraag of en in hoeverre eenvoudige en op zichzelf weinig belastende onderzoeken zoals het bepalen van het Hb daarbij in een bepaalde richting kunnen wijzen, wacht nog op antwoord.

Bij het zoeken naar dat antwoord is het van belang zich te realiseren dat voor de huisarts klachtgericht onderzoek meer aansluit op de probleemstellingen vanuit de dagelijkse praktijk dan ziektebeeldgericht onderzoek.

Zelfs als in de algemene populatie geen verband bestaat tussen Hb-gehalte en moeheid, kan dit verband nog wel aanwezig zijn bij patiënten die met de klacht moeheid op het spreekuur van de huisarts komen. Door ons wordt een onderzoek ingesteld met als vraagstelling: vindt de huisarts bij de mensen die hem of haar bezoeken wegens de klacht moeheid een andere verdeling van het Hb-gehalte dan bij overigens vergelijkbare spreekuurpatiënten zonder deze klacht?

De onderzoeksgroep bestaat uit patiënten die de huisarts voor de eerste keer bezoeken met de hoofdklacht moeheid. De controlegroep bestaat om redenen van praktische uitvoerbaarheid uit de op de 'cases' volgende patiënten zonder deze klacht, in dezelfde leeftijdscategorie en van hetzelfde geslacht. De klacht moeheid wordt daarbij zo gedefiniëerd, dat er voor de huisarts reden is om eventueel aan een verlaagd Hb te denken. Zo zal direct voor de hand liggende moeheid - bijvoorbeeld ten gevolge van een week slecht slapen na echtscheiding of na meer dan gebruikelijke inspanning - moeten worden onderscheiden van moeheid 'e causa ignota'. Ook zal een schaal naar graden van moeheid gehanteerd worden. In de controlegroep mogen geen personen voorkomen met klachten als duizeligheid of hartkloppingen, die naast moeheid geboekt staan als tekenen van een laag Hb.

## Summary

There is incomplete knowledge about the significance of the measurement of the blood haemoglobin level, and the relevance of this level to health. In

a retrospective study the haemoglobin levels were compared, of patients examined by the general practitioner because of fatigue (n=566) and a controlgroup of subjects without a special complaint having a general check-up (n=112). No relevant differences were found. Only for older men the haemoglobin level was somewhat lower in the group of patients with fatigue. Literature suggests little association of the 'classic symptoms' of anemia and haemoglobin level. It is not clear whether in general practice different distributions of haemoglobin levels are found in subjects with 'anemic' complaints such as fatigue and subjects without these complaints. A comparative study on the relation of the haemoglobin level and the frequently presented symptom of fatigue is planned.

## Literatuur

- Berry, W. T. C. & F. A. Nash, Symptoms as a guide to anaemia. *Brit. Med. J.* 1 (1954) 918
- Dawson, A. A., D. Ogston & H. W. Fullerton, Evaluation of diagnostic significance of certain symptoms and physical signs in anaemic patients. *Brit. Med. J.* 3 (1969) 436-439
- Edgerton, V. R., G. W. Gardner, Y. Ohira, K. A. Gunawardena & B. Senewiratne, Iron-deficiency anaemia and its effect on worker productivity and activity patterns. *Brit. Med. J.* 2 (1979) 1546-1549
- Elwood, P. C. & M. M. Wood, Effect of oral iron therapy on the symptoms of anaemia. *Brit. J. Prev. Soc. Med.* 20 (1966) 172-175
- Elwood, P. C., W. E. Waters, W. J. W. Greene, P. Sweetnam & M. M. Wood, Symptoms and circulating haemoglobin level. *J. Chron. Dis.* 21 (1969) 615-628
- Fry, J., Clinical patterns and course of anaemia in general practice. *Brit. Med. J.* 2 (1961) 1732-1736
- Garby, L., The normal haemoglobin level. *Brit. J. Haematol.* 19 (1970) 429-434
- Grol, R. (red), Huisarts en somatische fixatie. Nijmeegs Universitair Huisartsen Instituut, Nijmegen 1981
- Kelly, A. & L. Munan, Haematologic profile of natural populations: red cell parameters. *Brit. J. Haematol.* 35 (1977) 153-160
- Kirschenfeld, J. J. & H. H. Tew, Prevalence and significance of anaemia as seen in a rural general practice. *JAMA* 158 (1955) 807-811
- Lamberts, H., Redenen om naar de huisarts te gaan. *Huisarts & Wetensch.* 25 (1982) 301-310
- Ohira, Y., V. R. Edgerton, G. W. Gardner, K. A. Gunawardena & B. Senewiratne, S. Ikawa, Work capacity after iron treatment as a function of hemoglobin and iron deficiency. *J. Nutr. Sci. Vitaminol.* 27 (1981) 87-96
- Viteri, F. E., B. Torón, Anaemia and physical work capacity. *Clinics of Haematology* 3 (1974) 609-626
- Waters, W. E. & P. C. Elwood, Anaemia. In: D. L. Miller & R. D. T. Farmer (eds), *Epidemiology of diseases*, pp. 152-159. Blackwell, Oxford 1982
- Watt, C. M., The estimation of haemoglobin and the diagnosis of anaemia in general practice. *Practitioner* 209 (1972) 218-222
- Wood, M. M. & P. C. Elwood, Symptoms of iron deficiency anaemia, a community survey. *Brit. J. Prev. Soc. Med.* 20 (1966) 117-121
- Zoll, P. M., S. Wessler & M. J. Schlesinger, Interarterial coronary anastomoses in the human heart, with particular reference to anemia and relative cardiac anoxia. *Circulation* 14 (1951) 797-815

## Correspondentieadres

A. Knottnerus, arts, Capaciteitsgroep Epidemiologie, RU Limburg, postbus 616, 6200 MD Maastricht, tel. 043-888142

## Hoofdstuk 6

### Onverklaarde moeheid en hemoglobinegehalte; een onderzoek vanuit de huisartsenpraktijk

J. A. KNOTTNERUS, P. G. KNIPSCHILD, J. W. J. VAN WERSCH EN A. H. J. SIJSTERMANS

Moeheid is een klacht waarmee de huisarts vaak te maken krijgt.<sup>1</sup> Deze wordt daarmee voor een moeilijk diagnostisch probleem gesteld. Allerlei psychische of sociale problemen kunnen de oorzaak zijn, maar ook een lichamelijke ziekte.<sup>2,3</sup> In het laatste geval zou uitvoeriger diagnostiek meer aanknopingspunten kunnen bieden, terwijl deze in het eerste geval juist minder gewenst kan zijn wegens de kans op somatische fixatie.<sup>4</sup> Als anamnese en lichamelijk onderzoek onvoldoende aanknopingspunten bieden en als een psychosociale oorzaak niet voor de hand ligt, kan de diagnose bloedarmoede worden overwogen. Hematologisch onderzoek vindt dan vaak plaats, vaak als onderdeel van een batterij van tests. Het is echter nog maar de vraag of moeheid als klacht die geuit wordt tegenover de huisarts, een relevante indicatie is om dergelijk onderzoek te doen.

Uit de huisartsenpraktijk zijn twee onderzoeken gepubliceerd waarin het hemoglobine (Hb)-gehalte van patiënten met klachten die aan bloedarmoede doen denken, werd vergeleken met dat van een controlegroep zonder deze klachten. In het eerste onderzoek telde de controlegroep echter een aanzienlijk aantal zwangeren,<sup>5</sup> wat een eventueel verschil kan hebben gemaskeerd. Het tweede onderzoek had slechts betrekking op patiënten die naar een diagnostisch centrum verwezen waren.<sup>6</sup>

Het hier te beschrijven onderzoek betreft patiënten met 'onverklaarde moeheid'. De huisarts had voor hun moeheid bij het eerste consult geen voor de hand liggende verklaring kunnen vinden en achtte daarom een Hb-bepaling geïndiceerd.

#### PATIËNTEN EN METHODEN

Het onderzoek werd gedurende het eerste half jaar van 1984 verricht, in samenwerking met negen huisartsen te

#### SAMENVATTING

Bij 174 patiënten met onverklaarde moeheid bij wie de huisarts bepaling van het Hb-gehalte geïndiceerd achtte en bij 148 controlepatiënten bij wie dit niet het geval was, werd op gestandaardiseerde wijze hematologisch onderzoek verricht. In beide groepen werd dezelfde verdeling van het Hb-gehalte gevonden. Ook de andere hematologische waarden vertoonden geen verband met moeheid. Hb-waarden lager dan 7,0 mmol/l kwamen niet voor. Als in de huisartsenpraktijk bloedarmoede wordt vermoed alleen op grond van de klacht moeheid, dan zal dit vermoeden slechts zelden worden bevestigd.

Heerlen met een gezamenlijke praktijkomvang van 24.850 personen. Het onderzoek werd beperkt tot personen van 18 jaar en ouder, met uitsluiting van zwangeren en van patiënten die in de twaalf maanden vóór het consult ijzertherapie hadden gehad.

In de groep patiënten met onverklaarde moeheid werden al degenen opgenomen die met moeheid als nieuwe klacht op het spreekuur verschenen en die het spreekuur in de twaalf maanden daarvoor niet wegens deze klacht hadden bezocht. Als eis werd verder gesteld dat anamnese en lichamelijk onderzoek tijdens het eerste consult geen voor de hand liggende verklaring voor de klacht hadden opgeleverd, zoals duidelijke psychosociale problemen, bijwerkingen van geneesmiddelen of reeds bekende organische ziekten. Bovendien gold als criterium dat de huisarts bepaling van het Hb-gehalte geïndiceerd achtte.

Aan patiënten die op het spreekuur na hen kwamen en die van hetzelfde geslacht en ongeveer dezelfde leeftijd als de 'moeheidspatiënten' waren, werd gevraagd als controlepersonen te fungeren. De controlepersonen hadden geen klachten of ziekten die met anemie in verband zouden kunnen staan en zij hadden de huisarts in de voorgaande twaalf maanden ook niet daarvoor geconsulteerd. Voor de controlepersonen gold verder als eis dat de huisarts een Hb-bepaling bij hen niet geïndiceerd achtte.

Aan beide groepen patiënten werd door de huisarts gevraagd naar moeheid, duizeligheid, hartkloppingen,

Rijksuniversiteit Limburg, Capaciteitsgroep Epidemiologie, Postbus 616, 6200 MD Maastricht.

J. A. Knottnerus; prof. dr. P. G. Knipschild, sociaal-geneeskundige.

De Wever Ziekenhuis, Hematologisch Laboratorium, Heerlen.

Dr. J. W. J. van Wersch, klinisch chemicus.

A. H. J. Sijstermans, huisarts te Heerlen.

Correspondentie-adres: J. A. Knottnerus.

TABEL 1. Leeftijd en geslacht van 322 patiënten met en zonder onverklaarde moeheid

leeftijd (jaren)	aantal mannen		aantal vrouwen	
	moeheid	controlegroep	moeheid	controlegroep
18-39	18	25	89	46
40-59	18	16	25	33
60+	6	6	18	22
totaal aantal	42	47	132	101

TABEL 2. Hemoglobinegehalte bij 322 patiënten met en zonder moeheid

hemoglobine- gehalte (mmol/l)	aantal mannen (%)		aantal vrouwen (%)	
	met moeheid (n=42)	controle- personen (n=47)	met moeheid (n=132)	controle- personen (n=101)
7.0- 7.4	-	-	-	2 (2)
7.5- 7.9	1 (2)	1 (2)	9 (7)	5 (5)
8.0- 8.4	1 (2)	-	19 (14)	10 (10)
8.5- 8.9	4 (10)	2 (4)	42 (32)	37 (37)
9.0- 9.4	5 (12)	10 (21)	40 (30)	34 (34)
9.5- 9.9	11 (26)	19 (40)	20 (15)	8 (8)
10.0-10.4	11 (26)	11 (23)	1 (1)	4 (4)
10.5-10.9	4 (10)	3 (6)	1 (1)	-
11.0-11.4	5 (12)	1 (2)	-	1 (1)
gemiddeld Hb-gehalte* (standaard- afwijking)	9.9 (0.8)	9.7 (0.6)	8.9 (0.6)	8.9 (0.6)

\*normale waarden te Heerlen: mannen: 8.5-11.0 mmol/l, vrouwen: 7.5-10.0 mmol/l.

kortademigheid en pijn op de borst. In geval van moeheid werd naar de duur en de ernst ervan gevraagd. Voorts werd genoteerd of er sprake was van een bleek gelaat of nagelafwijkingen (longitudoonaal gegroefde nagels, brokkelige nagels, een bleek nagelbed, lepeltnagels). Tenslotte noteerde de huisarts telkens of hij een laag Hb-gehalte als verklaring van de klachten aanneemelijk vond.

Door de praktijkassistenten werd bij alle personen bloed afgenomen door venapunctie. Het bloed werd opgevangen in EDTA bevattende buizen en onderzocht in een centraal laboratorium (De Wever Ziekenhuis, Heerlen), doorgaans binnen enkele uren na afname. De hematologische waarden werden machinaal bepaald met een 'H-6000 analyzer' (Technicon Instruments).

## RESULTATEN

In totaal werden 322 patiënten volgens de criteria in het onderzoek betrokken. De tijdstippen waarop de beide groepen patiënten venapunctie ondergingen, waren gelijk over de dag verdeeld. De verdeling van de patiënten naar geslacht, leeftijd en categorie wordt gegeven in tabel 1. Bij vrouwen van jongere leeftijd waren er meer moeheidspatiënten dan controlepersonen. Dit werd niet veroorzaakt door weigeringen van patiënten, maar doordat het in de praktijk niet altijd lukte om consequent het één-op-één-schema aan te houden bij het vragen van controlepersonen na patiënten met moeheid als hoofdklacht. Er is geen reden om aan te nemen dat de uitkomsten van ons onderzoek hierdoor zijn beïnvloed.

In de tabellen wordt telkens onderscheid gemaakt naar geslacht, omdat het bekend is dat vrouwen een lager Hb-gehalte hebben dan mannen. Voor de eenvoud is geen onderscheid gemaakt naar leeftijd, omdat uit ons onderzoek bleek dat leeftijd geen invloed had op de bestudeerde verbanden. (In de onderzochte leeftidsgroep werd ook geen verband tussen Hb-gehalte en leeftijd gevonden.) De verdelingen en de gemiddelden van de Hb-concentraties bleken voor de moeheidsgroep en de controlegroep niet te verschillen (tabel 2). Als we de door het ziekenhuislaboratorium gehanteerde ondergrenzen van de referentiewaarden toepassen (mannen: 8.5 mmol/l en vrouwen: 7.5 mmol/l), dan vallen in de moeheidsgroep bij de mannen slechts twee uitslagen daaronder (5%) en bij de vrouwen geen enkele. Ook bij de mannen was er geen uitslag lager dan 7.5 mmol/l. Alleen in de controlegroep bij de vrouwen werden twee uitslagen lager dan 7.5 mmol/l gevonden.

De duur en de ernst van de klacht moeheid (beide gemeten op een driepuntsschaal) bleken niet van invloed te zijn. Het maakte ook niet uit of er behalve moeheid nog één of meer andere op anemie wijzende klachten of symptomen waren. Ook voor de andere hematologische waarden konden geen verschillen worden aangetoond tussen patiënten met moeheid als klacht en controlepersonen (tabel 3).

In de subgroep moeheidspatiënten bij wie volgens de huisarts een laag Hb-gehalte aanneemelijk was, werd alleen bij de vrouwen een iets lager gemiddeld Hb-gehalte gevonden (tabel 4). Meer betekenis zou gehecht kunnen worden aan de bevinding dat alleen in de groep vrouwen bij wie anemie aanneemelijk werd geacht, een aantal vrouwen voorkwam met een Hb-gehalte lager dan

TABEL 3. Gemiddelde (en standaardafwijking) van hematologische bepalingen bij 322 patiënten met en zonder moeheid als klacht

bepalingen	gemiddelde (standaardafwijking)			
	mannen		vrouwen	
	met moeheid (n=42)	controlepersonen (n=47)	met moeheid (n=132)	controlepersonen (n=101)
hemoglobinegehalte (mmol/l)	9.9 (0.8)	9.7 (0.6)	8.9 (0.6)	8.9 (0.6)
hematocriet (%)	47 (4)	47 (3)	43 (3)	43 (3)
aantal erythrocyten ( $\times 10^{12}/l$ )	5.1 (0.4)	5.1 (0.4)	4.6 (0.3)	4.7 (0.3)
MCV (fl)	93 (6)	92 (3)	93 (4)	92 (4)
MCH (fmol)	1.94(0.14)	1.91(0.08)	1.92(0.09)	1.90(0.10)
MCHC (mmol/l)	20.9 (0.5)	20.7 (0.4)	20.7 (0.4)	20.7 (0.4)



TABEL 4. Verdeling van het hemoglobinegehalte bij 174 patiënten met onverklaarde moeheid, naar aanneemelijkheid van een laag hemoglobinegehalte

hemoglobinegehalte (mmol/l)	laag hemoglobinegehalte			
	mannen (%)		vrouwen (%)	
	aanneemelijk	niet aanneemelijk	aanneemelijk	niet aanneemelijk
7,0- 7,4	-	-	-	-
7,5- 7,9	-	1 (4)	9 (11)	-
8,0- 8,4	-	1 (4)	14 (17)	5 (10)
8,5- 8,9	2 (13)	2 (8)	23 (27)	19 (40)
9,0- 9,4	2 (13)	3 (12)	27 (32)	13 (27)
9,5- 9,9	3 (19)	8 (31)	9 (11)	11 (23)
10,0-10,4	5 (31)	6 (23)	1 (1)	-
10,5-10,9	2 (13)	2 (8)	1 (1)	-
11+	2 (13)	3 (12)	-	-
<b>totaal</b>	<b>16 (100)</b>	<b>26 (100)</b>	<b>84 (100)</b>	<b>48 (100)</b>
<b>gemiddeld Hb-gehalte (standaardafwijking)</b>	<b>10,0 (0,8)</b>	<b>9,8 (0,9)</b>	<b>8,8 (0,6)</b>	<b>9,0 (0,5)</b>

8,0 mmol/l (11%). Waarden lager dan 7,5 mmol/l werden ook hier echter niet gevonden. Beide mannen met een Hb-gehalte onder de 8,5 mmol/l behoorden tot de subgroep moeheidspatiënten bij wie een laag Hb-gehalte niet aanneemelijk werd geacht (zie tabel 4).

Het verband tussen de klacht moeheid en de hematologische uitkomsten werd tenslotte nog onderzocht met multiële lineaire regressie-analyse. De reeds beschreven bevindingen werden hierdoor bevestigd.

#### BESCHOUWING

Patiënten consulteren hun huisarts in het algemeen wegens klachten en symptomen. Als anamnese en lichamelijk onderzoek weinig gegevens opleveren, moet de huisarts beslissen of nader onderzoek noodzakelijk is, bijvoorbeeld hematologisch onderzoek ter opsporing van bloedarmoede. Ons onderzoek geeft echter geen aanwijzingen voor een verband tussen onverklaarde moeheid en uitslagen van hematologische bepalingen zoals van het Hb-gehalte. De voorspellende waarde van de klacht moeheid voor de diagnose bloedarmoede lijkt gering te zijn. Het onderscheidend vermogen wordt er niet beter op; als tevens de duur en ernst van de klacht of andere 'anemische' klachten of verschijnselen in de beoordeling worden betrokken, of als de huisarts een laag Hb-gehalte (extra) aanneemelijk acht.

Het ontbreken van een verband tussen moeheid en Hb-concentratie komt ook naar voren uit onderzoek buiten de huisartsenpraktijk, in de algemene bevolking,<sup>7,9</sup> de polikliniek<sup>10</sup> en het ziekenhuis.<sup>11</sup> Dergelijke onderzoeken zijn voor de huisarts echter minder interessant en bovendien was daarin de mogelijke 'ruis' als gevolg van voor de hand liggende andere verklaringen voor moeheid niet uitgesloten.

Onze resultaten betekenen niet dat zéér lage Hb-waarden geen moeheidsklachten kunnen veroorzaken. In het gebied onder 6,5 mmol/l lijkt er wel verband te bestaan tussen het Hb-gehalte en bepaalde klachten zoals moeheid.<sup>12</sup> Extreem lage Hb-waarden worden echter in de huisartsenpraktijk maar weinig aangetroffen. In ons

onderzoek werd geen enkele maal een Hb-gehalte onder de 7,0 mmol/l gevonden.

Bij patiënten met klachten over moeheid die niet eenvoudig verklaard kan worden door psychosociale of andere oorzaken, denkt de huisarts al snel aan bloedarmoede. Vaak wordt dan hematologisch onderzoek verricht. Gezien onze bevindingen zal het geen verbazing meer wekken dat daarbij het vermoeden van bloedarmoede slechts zelden wordt bevestigd.

Wij danken P.G.J. van Aubel, J.P.H. Dolhain, P. Höpener, H.W.M. Logister, A.C.J. Schlösser, G.H.A. Siemons, A.H.J. Sijstermans, B.T.M.G. van der Werf, F.L.M. Zwanikken, huisartsen te Heerlen, en hun praktijkassistenten voor de patiëntengegevens, en A. Vissers, research-assistent van de Capaciteitsgroep Epidemiologie, die in alle fasen van het onderzoek nauw bij de uitvoering betrokken was.

#### SUMMARY

*Unexplained fatigue and haemoglobin concentration; a study in general practice patients.* - Standardized haematological tests were carried out in 174 patients with unexplained fatigue whose GPs considered determination of the Hb concentration indicated and in 148 controls without such fatigue. In both groups, the same distribution of haemoglobin values was found. Other haematological parameters showed no correlation with fatigue, either. Hb levels below 7.0 mmol/l were not found. If in general practice anaemia is suspected solely on the basis of the symptom of fatigue, this suspicion will be confirmed only rarely.

#### LITERATUUR

- Lamberts H. Redenen om naar de huisarts te gaan. *Huisarts Wet* 1982; 25: 301-10.
- Katerndahl DA. Fatigue of uncertain etiology. *Family Medicine Review* 1983; 1: 26-38.
- Morrison JD. Fatigue as a presenting complaint in family practice. *J Fam Pract* 1980; 10: 795-801.
- Grol R, red. Huisarts en somatische fixatie. Nijmegen: Nijmeegs Universitair Huisartsen Instituut, 1981.
- Watt CM. The estimation of haemoglobin and the diagnosis of anaemia in general practice. *Practitioner* 1972; 209: 218-22.

- <sup>6</sup> Knottnerus JA, Knipschild PG. Hemoglobine en 'anemische' klachten zoals moeheid. Tijdschr Soc Gezondheidszorg 1983; 61: 788-92.
- <sup>7</sup> Berry WTC, Nash FA. Symptoms as guide to anaemia. Br Med J 1954; i: 918.
- <sup>8</sup> Wood MM, Elwood PC. Symptoms of iron deficiency anaemia; a community survey. Br J Prevent Soc Med 1966; 20: 117-21.
- <sup>9</sup> Elwood PC, Waters WE, Greene WJW, Sweetnam P, Wood MM. Symptoms and circulating haemoglobin level. J Chronic Dis 1969; 21: 615-28.
- <sup>10</sup> Judy HE, Price NB. Haemoglobin level and red blood cell count findings in normal women. JAMA 1958; 167: 563-6.
- <sup>11</sup> Dawson AA, Ogston D, Fullerton HW. Evaluation of diagnostic significance of certain symptoms and physical signs in anaemic patients. Br Med J 1969; iii: 436-9.
- <sup>12</sup> Hjelm M, Wadman B. Clinical symptoms, haemoglobin concentration and erythrocyte biochemistry. Clin Haematol 1974; 3: 689-703.



---

SERUMFERRITINE, BIJ MOEHEIDSPATIËNTEN EN GEZONDE CONTROLEPERSONEN IN  
DE HUISARTSPRAKTIJK.

---

J.A. Knottnerus <sup>1)</sup> - P.G. Knipschild <sup>1)</sup> - J.W.J. van Wersch <sup>2)</sup> -  
A.H.J. Sijstermanns <sup>3)</sup>

Samenvatting.

Bij 174 volwassen patiënten met moeheidsklachten die de huisarts bezochten, en bij 148 controlepersonen werd naast de haematologische parameters het serumferritinegehalte bepaald. Er werden geen belangrijke verschillen tussen deze groepen gevonden. Onafhankelijk van de moeheidsklachten werd bij 11% van de onderzochten een serumferritinegehalte van 12 µg/l of lager gemeten. Het serumferritinegehalte bleek toe te nemen met de leeftijd (1.4 µg/l per jaar) en lager te zijn bij vrouwen dan bij mannen.

Tussen de hematologische parameters enerzijds en het serumferritine anderzijds werden geen significante verbanden gevonden. Moeheidsklachten en het hemoglobinegehalte lijken van beperkte waarde te zijn voor het stellen van de indicatie voor een serumferritinebepaling.

Trefwoorden: serumferritine/moeheid/huisartspraktijk/hematologie/parameters/

- 1) Capaciteitsgroep Epidemiologie, Rijksuniversiteit Limburg.
- 2) Haematologisch Laboratorium, Ziekenhuis de Wever, Heerlen.
- 3) Huisarts te Heerlen.

## Inleiding.

In een vorige publicatie berichtten wij geen verband gevonden te hebben tussen onverklaarde moeheidsklachten, gepresenteerd bij de huisarts, en het hemoglobinegehalte (Hb) (1). Omdat het echter denkbaar is dat de klachten wél samenhangen met een ijzertekort dat zich niet of nog niet heeft geuit in een verlaging van het Hb (2, 3), bestudeerden wij ook dit aspect.

Als maat voor de ijzervoorraad kozen wij het serumferritine. De beste informatie over de ijzerreserve krijgt men weliswaar door beenmerg-onderzoek, maar in de huisartspraktijk beperkt men zich bij de evaluatie van de ijzerstatus meestal tot serumbepalingen. Als de meest betrouwbare hiervan wordt tegenwoordig het serumferritine beschouwd. Het zou een gevoelige parameter zijn voor de ijzervoorraad van gezonde personen en personen met een ijzergebreksanemie (4-12). Het serumferritine zou al dalen, vóórdat een daling in serumijzer of hemoglobine wordt waargenomen. Aangenomen wordt dat een serumferritine concentratie lager dan 10 à 15  $\mu\text{g/l}$  altijd op ijzergebrek wijst (5). Vaak wordt een ondergrens van 12  $\mu\text{g/l}$  aangehouden (7, 8).

Behalve in het verband tussen moeheidsklachten en het serumferritine gehalte waren wij ook geïnteresseerd in de relatie tussen hematologische parameters en het serumferritine. In de praktijk wordt de indicatie tot het bepalen van het serumferritine immers mede gesteld op geleide van hematologisch onderzoek.

## Patiënten en methoden.

Het onderzoek vond plaats in het kader van het onderzoek naar het verband tussen hemoglobine en onverklaarde moeheid, waarover reeds gerapporteerd is (1). Negen Heerlense huisartsen selecteerden de patiënten gedurende het eerste half jaar van 1984. Uitgesloten werden zwangeren en personen jonger dan 18 jaar.

In de groep moeheidspatiënten werden alle patiënten opgenomen die de huisarts bezochten vanwege de klacht 'moeheid', en daarmee in de voorafgaande 12 maanden niet gezien waren. Uitgesloten werden personen die in die periode ijzertherapie hadden gehad en patiënten bij wie de moeheid in de loop van het consult te verklaren was door psychosociale

achtergronden, geneesmiddelengebruik of reeds bekende organische pathologie. Een Hb-bepaling werd in deze groep geïndiceerd geacht.

Als controlepersonen werden gevraagd op de moeheidspatiënten volgende spreekuurpatiënten, die van hetzelfde geslacht en ongeveer even oud waren en geen klachten hadden die in verband zouden kunnen staan met bloedarmoede, noch organische pathologie. Bij deze personen werd een Hb door de huisarts niet geïndiceerd geacht.

Van alle moeheidspatiënten en controlepersonen werd in zittende positie door middel van venapunctie bloed afgenomen, en opgevangen in EDTA-vacuumbuizen (Becton & Dickinson). Daarnaast werd een stolbuis afgenomen voor de bepaling van serumferritine.

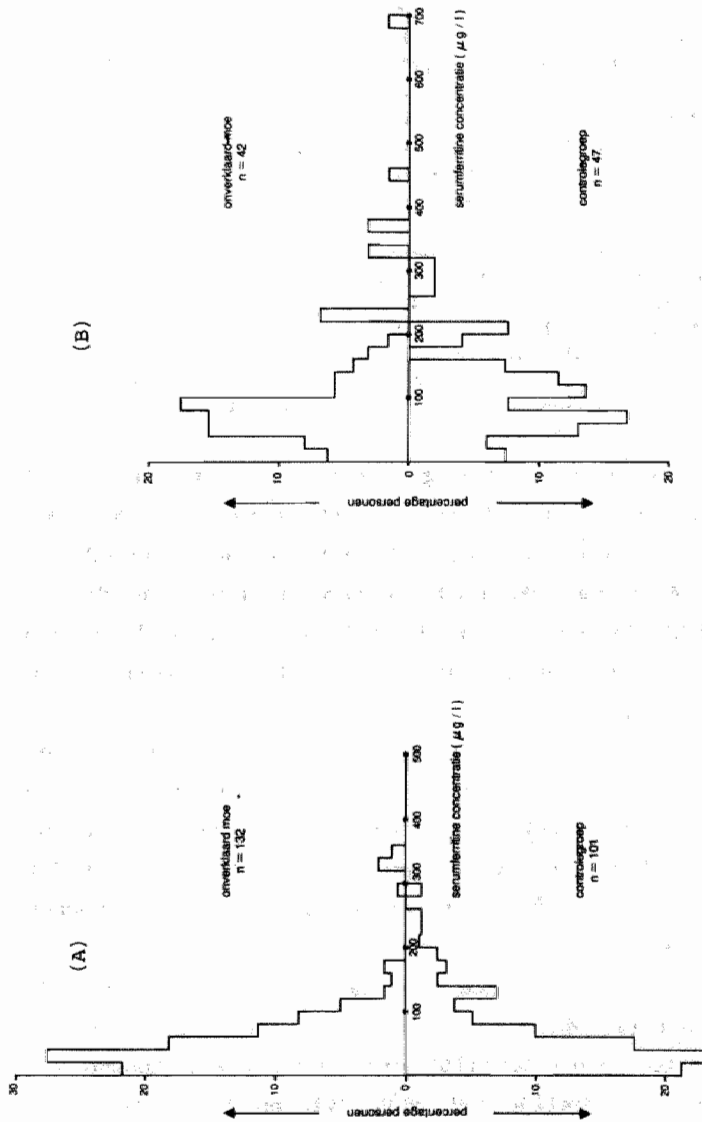
De hematologische parameters, werden enkele uren na bloedafname machinaal bepaald (H-6000, Technicon Instruments). Serum werd opgespaard en bij  $-70^{\circ}\text{C}$  ingevroren. De ferritinebepalingen werden later in serie verricht met een radio immuno assay techniek met behulp van reagentia van Becton & Dickinson.

De serumferritineconcentraties van beide groepen werden vergeleken na standaardisatie van hun leeftijdsverdeling op die van de Nederlandse bevolking per 1 januari 1984. Ook werd nagegaan welk gedeelte van de uitslagen per groep  $12\text{ }\mu\text{g/l}$  of lager was. Vervolgens werd het verband tussen serumferritine, leeftijd, geslacht en hematologische parameters bestudeerd. Er werd een multipele lineaire regressie-analyse verricht met de serumferritineconcentratie als afhankelijke variabele en de hematologische parameters, de leeftijd en het geslacht als onafhankelijke variabelen. Omdat de serumferritine distributie sterkt asymmetrisch was werd deze analyse herhaald na logaritmische transformatie van de serumferritineconcentratie.

### Resultaten.

Er werden 174 patiënten met onverklaarde moeheidsklachten (42 mannen en 132 vrouwen) en 148 controlepersonen (47 mannen en 101 vrouwen) in het onderzoek opgenomen.

In figuur 1 geven wij de voor de leeftijd gestandaardiseerde verdeling van het serumferritine voor vrouwen respectievelijk mannen, onderverdeeld naar moeheids- en controlepatiënten. Behoudens het voorkomen van wat meer hoge waarden bij de mannen in de moeheidsgroep, zijn er geen relevante verschillen tussen moeheids- en controlepersonen.



FIGUR 1: Distributie van serumferritine concentraties bij volwassen vrouwen (A) en mannen (B), gestandaardiseerd voor de leeftijd volgens de leeftijdsopbouw van de Nederlandse bevolking per 1 januari 1984 (histogram).

Tabel 1: Distributie van de serumferritine concentratie ( $\mu\text{g/l}$ ) naar leeftijd en geslacht. Range, kwartielen en percentages waarden van 12  $\mu\text{g/l}$  of lager.

mannen						
leeftijd	n	range	P25	P50	P75	% $\leq$ 12 $\mu\text{g/l}$
18-39	43	12-283	48	74	107	2,3
40-59	34	4-682	55	88	153	8,8
60 <sup>+</sup>	12	19-368	85	125	192	0,0
totaal	89	4-682	52	80	144	4,5

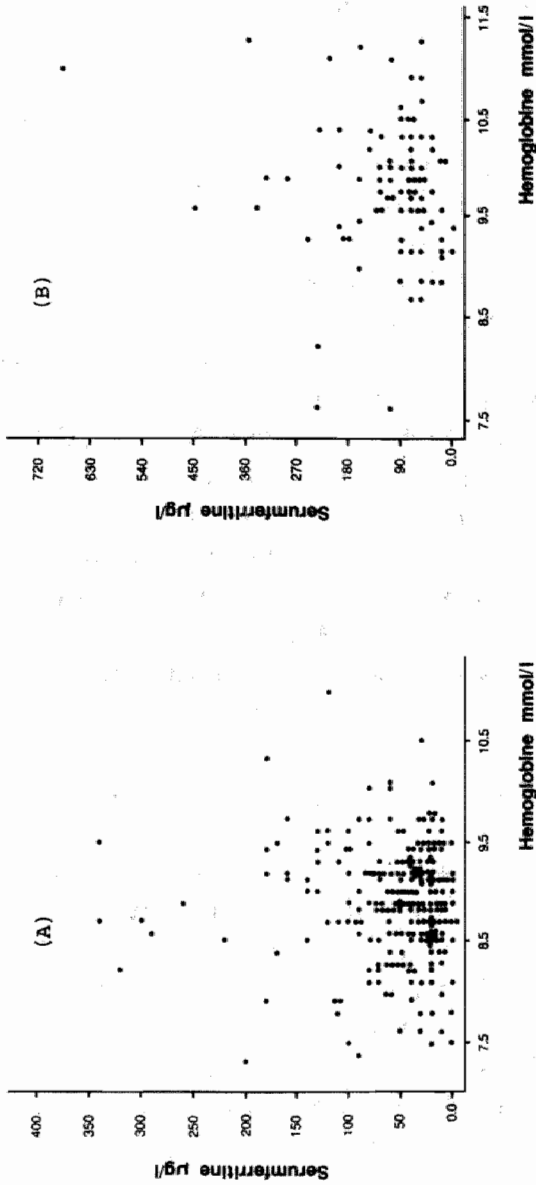
  

vrouwen						
leeftijd	n	range	P25	P50	P75	% $\leq$ 12 $\mu\text{g/l}$
18-39	135	4-287	17	31	55	14,8
40-59	58	6-324	24	46	76	10,3
60 <sup>+</sup>	40	4-344	49	89	140	10,9
totaal	233	4-344	21	41	74	13,3

In totaal was zowel in de moeheidsgroep als in de controlegroep 11% van de uitslagen lager dan of gelijk aan 12  $\mu\text{g/l}$ . Van de 42 mannen met moeheidsklachten had er één een serumferritine lager dan of gelijk aan 12  $\mu\text{g/l}$ , en van de 47 mannelijke controlepersonen drie. Van de 132 vrouwen met moeheidsklachten waren er 18 (14%) met waarden lager dan of gelijk aan 12  $\mu\text{g/l}$ , en van de 101 vrouwelijke controlepersonen waren dit er 13 (13%).

Omdat de moeheidsgroep en de controlegroep niet verschilden wat betreft de hematologische variabelen (1) en het serumferritinegehalte presenteren we de verdere beschrijving voor beide groepen tezamen. Tabel 1 geeft de verdeling naar leeftijd en geslacht. Men kan zien dat de serumferritine-concentratie toeneemt met de leeftijd, en bij mannen hoger is dan bij vrouwen.

In figuur 2 wordt gedemonstreerd hoe zwak het verband is tussen de serumferritine-concentratie en het Hb, bij vrouwen respectievelijk mannen. Hetzelfde patroon werd gevonden voor de verbanden tussen het serumferritine en respectievelijk het MCV, MCH en MCHC. Een lage serumferritine-concentratie (onder de 10 à 15  $\mu\text{g/ml}$ ) komt voor zowel bij de hoge als bij de lagere waarden van de hematologische parameters. Uit de multiële regressie-analyse bleek dat alleen de leeftijd



FIGUUR 2: Het verband tussen de hemoglobineconcentraties en de serumferritineconcentratie bij (A) vrouwen ( $n = 233$ ,  $r = 0.02$ ,  $p = 0.39$ ) en (B) mannen ( $n = 89$ ,  $r = 0.15$ ,  $p = 0.08$ ).

en het geslacht een significante relatie met het serumferritinegehalte vertonen ( $P < 0.05$ ). Per jaar leeftijdstoename stijgt het serumferritine gehalte met  $1,4 \mu\text{g/l}$ , en het serumferritine is bij mannen ongeveer  $45 \mu\text{g/l}$  hoger dan bij vrouwen. Herhaling van de analyse met de (normaal verdeelde) logaritme van de serumferritine-concentratie als afhankelijke variabele leverde dezelfde conclusies op.

#### Discussie.

Al eerder bleek dat wij bij volwassenen die zodanige moeheidsklachten hebben dat de huisarts een Hb-bepaling geïndiceerd acht, geen lagere waarden van Hb, MCV, MCH en MCHC vonden, dan bij mensen zonder die klachten (1). Nu blijken ook de serumferritine-concentraties niet lager te zijn. Als men uitgaat van de 'spreekuurpopulatie' van de huisarts is er dus weinig aanleiding om bij moeheidspatiënten een duidelijk lager serumferritinegehalte te verwachten dan bij andere patiënten. Rekening houdend met leeftijd en geslacht, konden wij geen relevant verband aantonen tussen de hematologische parameters enerzijds en het serumferritine anderzijds. De spreiding van het serumferritine rond iedere Hb-, MCV-, MCH- en MCHC-waarde is erg groot. Een serumferritine onder de norm blijkt net zo goed te kunnen voorkomen bij personen met een hoog Hb als bij personen met lagere Hb-waarden. Bij kinderen werd dit patroon al eerder beschreven (13). De leeftijd (oudere mensen hebben een hogere serumferritine-concentratie) en het geslacht (mannen hebben een hogere concentratie dan vrouwen) blijken wel belangrijke determinanten te zijn, zoals ook beschreven is door anderen (14, 15).

Een beperking van ons onderzoek is dat Hb-waarden lager dan  $7 \text{ mmol/l}$  niet werden gevonden. Des te opmerkelijker is het, dat er vrij veel serumferritine waarden van  $12 \mu\text{g/l}$  of lager gevonden werden (11%). Nader onderzoek naar de variatie van serumferritinewaarden bij zeer lage Hb-gehalten is gewenst. Indien ook daarbij een grote spreiding van de serumferritine-concentratie wordt gevonden, ontvalt de basis aan de strategie om de indicatie tot het bepalen van het serumferritine te stellen naar aanleiding van het vinden van een laag Hb.

Op zichzelf zijn onze bevindingen niet in tegenspraak met de opvatting dat de concentratie van het serumferritine indicatief is voor de ijzervoorraden in het reticulo-endotheliale systeem. Dit is met name beschreven bij gezonde personen en bij patiënten met ijzergebrek anemie (6-8). Bij onderliggende ziekten, waarbij sprake is van weefselafbraak of ontsteking (parenchymateuze leverziekten, leukemie, chronische infecties, rheumatoïde arthritis, gemetastaseerde carcino-men) is het serumferritine geen valide index van de ijzervoorraad meer (6, 7, 16). Dit zet een volgend vraagteken bij de waarde van de bepaling voor de diagnostiek, waarbij men van tevoren nu juist niet weet of de patient gezond is, of een bepaalde aandoening heeft.

De bezwaren die kleven aan de betekenis van het serumijzer en de transferrine-saturatie bij de diagnostiek van ijzergebrek (gevoeligheid voor aspecifieke prikkels, dagschommeling) (17, 18), kunnen aangevuld worden met twijfels ten aanzien van het serumferritinegehalte. Het kan een waarde hebben boven de referentiegrens (12 µg/l), niet alleen bij gezonde personen, maar ook bij patiënten met ijzergebrek gepaard aan allerlei ernstige aandoeningen. Ook kunnen blijkens ons onderzoek lage serumferritinewaarden voorkomen bij niet-vermoeide, niet-anemische personen.

Het klachtenpatroon, het hemoglobinegehalte en ook andere hematologische parameters lijken al met al weinig houvast te bieden bij het stellen van de indicatie voor het bepalen van het serumferritine.

#### LITERATUUR.

1. Knottnerus JA, Knipschild PG, Van Wersch JWJ, Sijstermans AHJ. Onverklaarde moeheid en hemoglobinegehalte, een onderzoek vanuit de huisartspraktijk. Ned Tijdschr Geneesk 1986; 130: 402-405.
2. Jacobs A. Non-haematological effects of iron deficiency. Clinics in Haematology 1982; 11: 353-364.
3. Waters WE, Elwood PC. Anaemia. In: Miller DL, Farmer RDT (eds.). Epidemiology of diseases. Oxford, Blackwell 1982: 152-159.
4. Eijk HG van, Heul C van der. Enkele facetten van de ijzerstofwisseling. Ned Tijdschr Geneesk 1978; 122: 1759-1765.
5. Zuyderhoudt FMJ. Ferritine in serum. Ned Tijdschr Geneesk 1982; 126: 1733-1736.



6. Worwood M. Ferritin in human tissues and serum. *Clinics in Haematology* 1982; 11: 275-307.
7. Jacobs A, Miller F, Worwood M et al. Ferritin in the serum of normal subjects and patients with iron deficiency and iron overload. *British medical Journal* 1972; 4: 206-208.
8. Jacobs A, Worwood M. The clinical use of serum ferritin estimation. *British Journal of Haematology* 1975; 31: 1-3.
9. Jacobs A, Path FRC, Worwood M. ferritin in serum. Clinical and biomedical implications. *New England Journal of Medicine* 1975; 292: 951-956.
10. Mazza J, Barr RM, McDonald JED et al. Usefulness of the serumferritin-concentration in the detection of iron deficiency in a general hospital. *canadian Medical Association Journal* 1978; 119: 884-886.
11. Linpisan S, Thananykul O, Suwanraj C et al. Iron deficiency in a Northern Thai population: the effects of iron supplements studied by means of plasma ferritin estimations. *Ann Clin Bioch* 1984; 21: 268-274.
12. Harju E, Pakarinen A, Larmi T. A comparison between serumferritin concentration and the amount of bone marrow stainable iron. *Scand J Clin Lab Invest* 1984; 44: 555-556.
13. Hershko C, Bar-Or D, Gaziel Y et al. Diagnosis of iron deficiency anemia in a rural population of children. Relative usefulness of serumferritin, red cell protoporphyrin, red cell indices, and transferrin saturation determinations. *American Journal of Clinical Nutrition* 1981; 34: 1600-1610.
14. Casale G, Bonora C, Migliavacca A et al. Serumferritin and ageing. *Age and Ageing* 1981; 10: 119-122.
15. Rubin C, Wood PJ, Archer T, Rowe DJF. Changes in serumferritin and other 'acute phase' proteins following major surgery. *Ann Clin Bioch* 1984; 21: 290-294.
16. Rothwell RS, Davis P. Relationship between serumferritin, anemia, and disease activity in acute and chronic rheumatoid arthritis. *Rheumatology Int* 1981; 1: 65-67.
17. Cavill I. Diagnostic methods. *Clinics in Haematology* 1982; 11: 259-273.
18. Cook JD. Clinical evaluation of iron deficiency. *Seminars in Hematology* 1982; 19: 6-18.

HET HEMOGLOBINEGEHALTE VAN ZWANGEREN EN SUBJECTIEVE KLACHTEN.

J.A. Knottnerus 1) - L. Delgado 2) - P.G. Knipschild 2) - G.G.M. Essed 3) - F. Smits 3)

Samenvatting.

Verslag wordt gedaan van een onderzoek naar het verband tussen subjectieve klachten en hematologische parameters bij zwangeren. Door 494 zwangeren werd rond de 30e zwangerschapsweek een vragenlijst ingevuld. Direct aansluitend werd bloed afgenomen. Met name bij de volgende klachten werden vaker lage Hb- en/of Ht- waarden gevonden: Moeheid (erger dan voor de zwangerschap), hartkloppingen, kuitkrampen (vooral als ze 's morgens optreden), en duizeligheid. Reeds onder de afkappunten 7.0 mmol/l (Hb) respectievelijk 32 % (Ht) is er een toename van deze klachten waar te nemen. Het verband tussen klachten en een laag Hb is vooral duidelijk bij zwangeren met een MCV van 90 of hoger. Slechts 4 % van alle zwangeren heeft geen enkele van de bestudeerde klachten, en ook bij vrouwen met hogere Hb- en Ht- waarden worden vrij vaak klachten gevonden. Door middel van een gemodificeerde stepwise multipele logistische regressie analyse werd nagegaan van welke klachten het aannemelijkheidsquotient ten aanzien van een Hb en/of Ht onder de genoemde afkappunten significant van 1.0 verschilde na correctie voor de aanwezigheid van eventuele andere klachten. Dit bleek te gelden voor moeheid (erger dan voor de zwangerschap) en hartkloppingen. De aannemelijkheidsquotienten werden niet wezenlijk beïnvloed door leeftijd, pariteit en het door een verloskundige of door een gynaecoloog gecontroleerd zijn. Het onderscheidend vermogen van de klachten lijkt evenwel te gering om het controleren van de hematologische parameters louter op geleide van bepaalde klachten plaats te doen vinden. Wel kunnen klachten aanleiding zijn voor een extra bloedcontrole. Nader onderzoek naar het verband tussen foliumzuurgebrek en subjectieve klachten bij zwangeren lijkt gewenst.

- 1) Capaciteitsgroep Epidemiologie en Capaciteitsgroep Huisartsgeneeskunde
  - 2) Capaciteitsgroep Epidemiologie
  - 3) Capaciteitsgroep Verloskunde en Gynaecologie
- Rijksuniversiteit Limburg

## Inleiding.

Het melden van klachten als moeheid, duizeligheid en hartkloppingen tijdens de prenatale controle leidt vaak tot extra bloedonderzoek en kan de doorslag geven bij de beslissing om al dan niet over te gaan tot nadere diagnostiek en/of behandeling. Wij konden echter geen onderzoekspublicatie vinden over het verband tussen subjectieve klachten en een laag Hb bij zwangeren. Het is nog maar de vraag, in hoeverre subjectieve klachten inderdaad de doorslag zouden moeten geven ten aanzien van diagnostiek en/of behandeling van anemie. In het kader van de bestudering van het verband tussen moeheidsklachten en het Hb in de huisartspraktijk werden zwangere vrouwen van het onderzoek uitgesloten (1). De redenen daarvoor waren dat, ten gevolge van hemodilutie het Hb vanaf ongeveer de 6e zwangerschapsweek met ongeveer 1 mmol/l daalt, en dat in de onderzoeksperiode te weinig zwangeren in een vergelijkbare fase van de graviditeit zich bij de deelnemende huisartsen zouden melden. Daarom besloten wij tot een speciaal onderzoek bij zwangeren. Gekozen werd voor een benadering van het referentiewaarden vraagstuk, waarbij werden bestudeerd zowel de relatie tussen het Hb en de hematocriet (Ht) en subjectieve klachten, als de relatie tussen het Hb en de Ht en het verloop en de uitkomst van de zwangerschap.

De resultaten van het onderzoek naar het verband tussen de hematologische parameters en de zwangerschapsuitkomst worden apart beschreven in hoofdstuk 22. In het nu volgende verslag staat de vraag centraal, of gedurende de zwangerschap bepaalde klachten indicatief zijn voor lage Hb- en/of Ht- concentraties, en, indien dergelijke klachten worden gevonden, onder welke Hb- en Ht- concentraties deze klachten vaker worden gevonden.

Bij het beantwoorden van deze vraag kozen wij voor het bestuderen van het verband tussen Hb en Ht en klachten bij de routinecontrole die rond de dertigste zwangerschapsweek plaatsvindt. Juist bij die controle zou vaak worden besloten over eventuele ijzersuppletie, en bovendien vindt er na de 32e week weer een geringe spontane stijging plaats (2,3).

### Methode.

In de periode van 1 juli 1985 tot en met eind april 1986 werd een vragenlijst voorgelegd aan de zwangeren die de polikliniek verloskunde van het Academisch Ziekenhuis Maastricht bezochten in verband met de eerste zwangerschapscontrole na de dertigste week. Gevraagd werd ondermeer naar:

- klachten (die met een laag Hb zouden kunnen samenhangen), waarvan men de voorafgaande week eventueel last had gehad (moeheid, hoofdpijn, duizeligheid, hartkloppingen, kortademigheid, bleekheid, concentratieproblemen, prikkelbaarheid, slaapstoornissen, neerslachtigheid, kramp in de kuiten). Over de moeheidsklachten werd gevraagd of deze reeds bij het opstaan aanwezig waren, of ze in de loop van de dag erger werden, en of men er af en toe dan wel dagelijks last van had. Ook over een aantal andere klachten werd meer specifieke informatie gevraagd. Ook werd gevraagd of men zelf een verklaring voor de klachten had, en zo ja welke. Door de onderzoekers werden deze verklaringen, zonder voorkennis van het Hb-gehalte, later ingedeeld in waarschijnlijk somatische, waarschijnlijk psychosociale en andere verklaringen. Gevraagd werd ook of men nu meer last van de klachten had dan vóór de zwangerschap.
- het gebruik van staaltabletten.

De vragenlijst werd vóór het verlaten van de polikliniek ingeleverd. Direct aansluitend werd bloed afgenomen door middel van venapunctie. Het bloed werd opgevangen in EDTA-buizen, en binnen enkele uren werden de hematologische parameters bepaald met een SYSMEX Microcell-counter (hematologisch laboratorium Academisch Ziekenhuis Maastricht, Mevrouw Dr. M.S.M. Daniels-Bosman en Dr. G.J.P.C.M. Kok).

De vragenlijst werd ook voorgelegd aan de zwangeren die begeleid werden in twee eerstelijns verloskundige praktijken, bij de eerste controle ná 30 weken amenorroe in de periode van 1 oktober 1985 tot 1 februari 1986. Direct aansluitend werd venapunctie verricht in de verloskundige praktijk of op één van de perifere prikpunten van de Prikdienst van het Academisch Ziekenhuis. Het bloed werd opgevangen in EDTA-buizen, en binnen 24 uur onderzocht in het hematologisch laboratorium van het Academisch Ziekenhuis Maastricht.

De gegevens werden bewerkt op de VAX-computer van de Rijksuniversiteit Limburg, en geanalyseerd met het BMDP-programma (4).

De statistische analyse werd begonnen met het bestuderen van kruistabellen aangaande de relatie klachten en hematologische parameters. Verschillen tussen gemiddelden werden getoetst met behulp van de t-toets voor ongepaarde waarnemingen, en verschillen tussen proporties met de chikwadraat toets, of de Fisher exact test indien het verwachte aantal patiënten in één van de cellen 5 of minder bedroeg. Vervolgens werd door middel van een forward stepwise multiple lineaire regressie-analyse nagegaan welke klachtenset significant samenhang met het Hb-gehalte. Door middel van een forward stepwise logistische regressie-analyse werd nagegaan welke klachten een sterke en significante samenhang vertoonden met een Hb-gehalte lager dan 7.0 mmol/l en/of een Ht van 32 % of lager. Een kenmerk van diagnostiek is dat zij uitgaat van de a priori kans op een bepaalde afwijking (welke vervolgens wordt gemodificeerd door de diagnostische gegevens) en niet van de a posteriori kans, gegeven een negatief resultaat van alle nog te stellen vragen of uit te voeren tests (welke laatstgenoemde kans in het algemeen wordt gepresenteerd in het intercept van de regressiefunctie). Daarom ontwikkelden wij een techniek waarbij het intercept van de logistische regressiefunctie de representatie wordt van de a priori kans in de bestudeerde populatie. Daardoor is het mogelijk rechtstreeks aannemelijkheidsquotiënten (likelihoodratio's, LR) te bepalen voor de positieve testuitslag of de aanwezigheid van een klacht of symptoom:

$$\frac{\text{sensitiviteit}}{(1-\text{specificiteit})}$$

en voor de negatieve testuitslag of de afwezigheid van een klacht of symptoom:

$$\frac{(1-\text{sensitiviteit})}{\text{specificiteit}}$$

gecorrigeerd voor de invloed van andere tests, symptomen of factoren (zie appendix).

Voorts werd bestudeerd, of de verbanden tussen klachten en hematolo-

gische parameters verschilden in de subpopulaties begeleid door respectievelijk eerstelijns verloskundigen, verloskundigen op de polikliniek en de gynaecologen. Tenslotte werd nagegaan of de verbanden anders waren bij vrouwen die al dan niet staaltabletten hadden gebruikt en bij vrouwen met een laag of hoger Mean Cellular Volume (MCV).

### Resultaten.

Van de zwangeren die via de polikliniek verloskunde benaderd werden, nam 92% deel aan het onderzoek. Redenen voor niet-deelname waren: weigering (4%), taalbarrière (2%), niet op de betreffende controle verschijnen (2%). Van de zwangeren die via de zelfstandig gevestigde verloskundigen benaderd werden, nam 63 % deel; 12 % weigerde, 10 % vond de afstand naar de prikdienst te groot, 8 % werd bij de controle gemist, en bij 7 % waren er diverse andere redenen.

Voor de analyse zijn beschikbaar de gegevens van 494 vrouwen (90 "perifeer" begeleid, en 404 poliklinisch). De leeftijdsrange was 18 tot 39 jaar (P50 = 27), en 50 % was nullipara. Bij alle deelnemende vrouwen is het Hb bepaald (gemiddeld 7.3 mmol/l range 5.6 - 9.1, SD: 0.6), bij 96% de hematocriet (gemiddeld 35%, range 26-44, SD: 3) en bij 91% het MCV (gemiddeld 91, range 75-114, SD: 5).

In tabel 1 is weergegeven welk percentage van de zwangeren per klasse van de drie hematologische parameters een bepaalde klacht meldde. Veel klachten bleken wat vaker voor te komen bij lagere hemoglobine- en hematocrietwaarden: moeheid, duizeligheid, hartkloppingen, kortademigheid, bleekheid en kuitkrampen. Hoofdpijn werd zowel in de hogere als in de lagere klassen vaker gevonden. De 'anemische' klachten moeheid, bleekheid en kortademigheid leken eerder geassocieerd te zijn met een hoger MCV dan met een lager MCV.

Over het geheel genomen komen de klachten veel voor, in iedere klasse van de hematologische parameters. Moeheid en kuitkrampen spannen wat dit betreft de kroon: respectievelijk 68 % en 50 % van de vrouwen bleek hiervan last te hebben gehad in de week voorafgaande aan het invullen van de vragenlijst. Slechts 4% van de vrouwen heeft géén enkele van de genoemde klachten.

Tabel 1: Percentage zwangeren met een bepaalde klacht, per Hb-, Ht- en MCV-klasse.

	HEMOGLOBINE (N=494)					HEMATOCRIET (N=474)					MCV (N=451)	
KLACHT	< 6.4 n=29	-6.9 n=107	-7.4 n=161	-7.9 n=136	> 80 n=61	< 30 n=42	-34 n=186	-38 n=196	> 39 n=50	<90 n=168	> 90 n=283	
MOEHEID	76	75	65	63	67	79	69	67	60	61	72	
HOOFDPIJN	21	11	13	13	21	21	15	10	22	13	13	
DUIZELIGHEID	35	24	19	17	18	33	24	18	14	21	19	
HARTKLOPPINGEN	21	20	11	10	10	21	15	13	4	12	14	
KORTADEMIGHEID	52	44	34	37	38	50	42	36	34	36		
BLEEKHEID	38	25	34	21	23	42	23	29	20	23	28	
CONCENTRATIE- MOEILIKHEDEN	10	10	11	17	10	14	8	15	14	8	15	
PRIKKELBAAKHEID	38	39	47	39	49	43	42	44	44	42	43	
SAPEN: slecht matig	21	11	16	22	15	19	12	18	18	13	16	
NIERSLAGTIGHEID	41	50	44	36	54	50	45	45	40	45	46	
KUITRAMPEN	18	20	21	18	25	22	18	24	14	21	20	
	63	54	46	50	45	63	53	47	40	51	48	

Tabel 2: Hematologische parameters per klacht. Gemiddelden, standaardafwijkingen (SD) en percentage waarden onder bepaalde afschappunten.

KLACHT	HEMOGLOBINE MMOL/L					HEMATOCRIET %				
	n	gemidd.	SD	% < 6,4	% < 6,9	n	gemidd.	SD	% < 30	% < 32
ALLEN (prior kans)	494	7,3	0,6	5,9	27,5	474	35	3	8,9	24,7
MOE	332	7,3	0,6	6,6	30,6*	320	34	3	10,2	27,3
al moe bij opstaan	116	7,3	0,6	8,6	30,9*	108	34+	3	13,9*	31,5
's middags erger	275	7,3	0,6	6,5	22,2	265	35	3	9,8	25,6
iedere dag	149	7,3	0,5	6,7	27,5	139	35	3	9,4	25,9
onverklaard	126	7,2+	0,6	8,7	34,9*	121	34*	3	12,4	30,6
erger dan voor zwangerschap	62	7,1+	0,6	16,1+	41,9+	58	34	3	17,2*	37,9*
HOOFDPIJN	68	7,4	0,7	8,8	26,4	67	35	4	13,4	19,4
iedere dag	8	7,7*	0,4			8	36	2		
onverklaard	30	7,3	0,6	6,7	30,0	29	35	3	10,3	16,9
erger dan voor zwangerschap	5	7,1	0,8			5	33	4		
DUIZELIG	101	7,2*	0,6	9,9	35,6*	100	34*	3	14,0*	34,0*
zwart voor de ogen	23	7,1	0,6	13,0	43,4	23	33*	3	26,0*	44,0*
iedere dag	11	7,0	0,6	18,2	36,4	11	32*	3	36,4*	45,6
erger dan voor zwangerschap	13	6,9*	0,5	15,4	53,9*	13	32+	3	30,8*	53,9*
HARTKLOPPINGEN	65	7,2*	0,6	9,2	41,5+	64	34*	3	14,1	36,0*
KORTADEMIG	189	7,3	0,6	7,9	32,2	185	34	3	10,8	27,6
BLEEK	133	7,2	0,5	8,3	27,8	126	34	3	13,5*	22,2
CONCENTRATIEPROBLEMEN	61	7,4	0,5	4,9	22,9	57	35	3	10,5	19,3
onverklaard	38	7,4	0,5	5,3	21,1	35	35	3	8,6	14,3
PIJKELBAAR	210	7,3	0,6	5,2	25,2	203	35	3	8,9	26,1
onverklaard	107	7,4	0,5	1,9	17,8	102	35	3	4,9*	18,6
SLAPEN: slecht	81	7,3	0,6	7,4	22,4	73	35	3	11,0	19,2
matig	216	7,3	0,6	5,6	30,1	213	35	3	9,9	27,7
TERNEERGESLAGEN	101	7,3	0,6	5,1	26,5	95	35	3	9,5	23,2
KUITKRAMPEN	215	7,3	0,6	7,9	31,2	207	34*	3	11,5	26,9
nacht	141	7,3	0,6	5,7	29,8	137	34	3	11,7	25,6
ochtend + nacht	47	7,1	0,6	17,0+	38,3	44	33	3	15,9	34,1
iedere dag	37	7,3	0,5	2,7	32,4	36	34	3	2,8	25,6

\*:  $p < 0,05$

+:  $p < 0,01$

De vergelijkingsgroepen bij de statistische toetsing waren de zwangeren met, respectievelijk zonder de betreffende klacht wat betreft het gemiddelde Hb en Ht en wat betreft de proportie uitslagen onder een bepaald afschappunt.



Omdat voor de arts subjectieve klachten als vertrekpunt minstens zo belangrijk zijn als laboratoriumuitslagen, is in tabel 2 voor elke klacht aangegeven, in welke mate deze een indicator is voor een laag Hb respectievelijk Ht. Per klacht is van bepaalde, tevens nagevraagde, nadere kenmerken ook het verband met Hb en Ht aangegeven. De klachten 'iedere dag hoofdpijn' en 'duizelig, erger dan vóór de zwangerschap' bleken gepaard te gaan met een duidelijk hoger, respectievelijk lager gemiddeld Hb dan dat van de gehele groep. Voor de gemiddelde hematocriet werd een dergelijk (duidelijk negatief) verband alleen gevonden voor duizeligheidsklachten, 'iedere dag', en 'erger dan vóór de zwangerschap'. Meer informatief is het beeld van de "voorspellende waarden" van de klachten ten aanzien van een laag Hb of Ht vanaf de afkappunten 6.4 en 6.9 mmol/l, respectievelijk 30 en 32 %. Ten behoeve van de duidelijkheid, en ook omdat dit op grond van de data te rechtvaardigen is, beperken wij ons in de verdere bespreking tot de afkappunten 6.9 mmol/l respectievelijk 32 %. In tabel 2 is ook de informatie voor de lagere afkappunten beschikbaar. Kijkend naar de afkappunten 6.9 mmol/l en 32 %, blijkt dat de "a posteriori" kansen op lage waarden (d.w.z. de kans hierop, gegeven bepaalde klachten) vooral toenemen ten opzichte van de "a priori kans" (d.w.z. de kans op lage waarden in de gehele groep) bij onverklaarde moeheid, "moeheid, erger dan vóór de zwangerschap", duizeligheid ("zwart voor de ogen", "iedere dag", "erger dan vóór de zwangerschap"), hartkloppingen, kuitkrampen ('s morgens, al dan niet gepaard aan nachtelijke krampen). Wanneer er sprake is van prikkelbaarheid waarvoor men zelf geen verklaring heeft, lijkt er juist minder vaak sprake van lage Hb- en Ht-waarden. Daartegenover staat, dat als men wél een verklaring heeft, het Hb relatief vaak lager dan 7.0 mmol/l is. Dat geldt vooral voor als "psychosociaal" geduide verklaringen (14/36, versus 19/107 in de "geen verklaring" groep,  $P < 0.05$ ).

In tabel 3 zijn de lineaire regressiecoëfficiënten gegeven van die klachten, die na selectie via een forward stepwise procedure een statistisch significante samenhang blijken te hebben met de hemoglobineconcentratie en/of de hematocriet. Deze invloeden lijken, behoudens voor duizeligheid, gering. Opmerkelijk is weer dat "prikkelbaarheid zonder verklaring" eerder met een hoog dan met een laag Hb/Ht samenhangt.

Tabel 3: Verband tussen hemoglobinegehalte respectievelijk hematocriet, en de klachtenset die geselecteerd werd door middel van een forward stepwise multiële lineaire regressie ( $p < 0,05$ ).

KLACHT	HEMOGLOBINE (mmol/l)		HEMATOCRIET (%)	
	aanwezig	niet aanwezig	aanwezig	niet aanwezig
MOEHEID, erger dan vóór de zwangerschap	-0,24	+0,02		
MOEHEID, die in de loop van de dag erger wordt	+0,03	-0,17		
MOEHEID, zonder verklaring	-0,08	+0,06		
HARTKLOPPINGEN			-0,08	+0,01
DUIZELIGHEID			-0,85	+0,21
KUITKRAMPEN			-0,38	+0,38
PRUKEELBAARHEID, zonder verklaring	+0,09	-0,12	+0,53	-0,72

Tabel 4: Aannemelijkheidsquotienten voor de aan- respectievelijk afwezigheid van bepaalde klachten geselecteerd en voor elkaars invloed gecorrigeerd door middel van stapwise logistische regressie (zie ook appendix).

KLACHT	(1)		(2)	
	aanwezigheid klacht	afwezigheid klacht	aanwezigheid klacht	afwezigheid klacht
moehheid, erger dan vóór zwangerschap *	2,0	0,89	2,1	0,88
duizeligheid	1,3	0,80		
hartkloppingen **	1,8	0,90	1,8	0,91
kuitkrampen ochtend + nacht	1,7	0,93		

model (1): toelatings- en eliminatiegrenzen 0,15  
model (2): toelatings- en eliminatiegrenzen 0,05

\* in model (1)  $p < 0,05$ , in model (2)  $p < 0,01$

\*\* in model (1)  $p < 0,05$ , in model (2)  $p < 0,05$

a priori kans op een  $Hb < 7,0$  mmol/l en/of een  $Ht < 32\%$  is 32%

posterior odds in geval van zowel moehheidsklachten als hartkloppingen, berekend uit model (2):

$0,32/0,68 \times 2,1 \times 1,8 = 1,76$ , overeenkomend met een posterior kans van 64%.

posterior odds in geval van noch moehheidsklachten, noch hartkloppingen, berekend uit model (2):

$0,32/0,68 \times 0,88 \times 0,91 = 0,37$ , overeenkomend met een posterior kans van 27%.

Tabel 4 presenteert de resultaten van de multiële logistische regressie, met afhankelijke variabele: de posterior odds van een Hb van 6.9 mmol/l of lager en/of een Ht van 32 % of lager. Gegeven zijn de aannemelijkheidsquotiënten van de klachten die geselecteerd werden door middel van een stepwise logistische regressie-analyse. Moeheidsklachten, die erger zijn dan vóór de zwangerschap, duizeligheid, hartkloppingen en kuitkrampen, die 's morgens + 's nachts optreden, blijken de kans op een Hb en/of Ht onder genoemde grenswaarden, ook los van elkaar, te verhogen. Deze klachtenset is het resultaat van zowel een forward als een backward stepwise procedure, met als toelatings- en eliminatiegrenzen p-waarden van 0.15. Indien we deze grenzen stringenter maken ( $p = 0.05$ ), blijven alleen de betreffende moeheidsklachten en hartkloppingen over als indicatoren van lage Hb- en/of Ht- waarden, met aannemelijkheidsquotiënten in dezelfde orde van grootte. Als pariteit (al dan niet nullipara) en leeftijd als onafhankelijke variabele in het model worden opgenomen, veranderen de uitkomsten niet. Het resultaat van deze multivariabele analyse impliceert, dat men de diagnostische informatie van de genoemde klachten aan elkaar kan toevoegen (zie appendix). Zo is op grond van de bepaalde waarden van de aannemelijkheidsquotiënten te berekenen, dat een zwangere die noch het genoemde type moeheidsklachten heeft noch hartkloppingen, nog een kans van 27 % heeft op een Hb en/of Ht onder de gestelde normen. Heeft men beide klachten wel dan wordt deze kans 64 % (zie toelichting tabel 4).

Om na te gaan of de bestudeerde relaties anders lagen in de populatie vrouwen die gedurende de gehele zwangerschap door een verloskundige (perifeer of poliklinisch,  $n=145$ ) begeleid was, dan in de populatie die primair of secundair door een gynaecoloog begeleid was ( $n=349$ ), werd voor beide subpopulaties een aparte regressie-analyse verricht. Moeheid, erger dan vóór de zwangerschap, en hartkloppingen bleken in beide gevallen weer de belangrijkste predictoren met aannemelijkheidsquotiënten die in dezelfde orde van grootte lagen als in de totale groep.

Omdat het op indicatie gebruikt hebben van staaltabletten (43% gebruikte deze vóór de 30e week) invloed kan hebben op het verband tussen klachten en hemoglobinegehalte, gingen wij na of de bestudeerde relaties veranderden als wij vrouwen die staal gebruikten hadden al dan

niet , c.q. apart, in de analyse betrokken. De richting van de verbanden veranderde niet en de coëfficiënten bleven weer in dezelfde orde van grootte. Het Hb van de staalgebruiksters was wel iets lager dan dat van de niet-staalgebruiksters (7.23 mmol/l (SD: 0.49) versus 7.34 mmol/l (SD: 0.61) ( $p < 0.05$ ).

Om een indruk te krijgen in hoeverre foliumzuurtekort een rol zou kunnen spelen, is het verband tussen klachten en Hb respectievelijk Ht apart bestudeerd voor vrouwen met een MCV lager dan, respectievelijk gelijk aan of groter dan 90 fl. Terwijl er bij lagere MCV- waarden vaker lage Hb- en Ht- concentraties voorkwamen dan bij hoge MCV- waarden, werden moeheid, hartkloppingen en kortademigheid relatief vaak aangetroffen in de groep met zowel een hoger MCV, als een lager Hb (tabel 5). Voor de hematocriet vonden we eenzelfde beeld. Bij de multipele lineaire en logistische regressie konden nu alleen verbanden tussen klachten (de klachtenset moeheid, erger dan vóór de zwangerschap, duizeligheid en kuitkrampen) en Hb- en/of Ht- waarden onder de gehanteerde afkappunten worden aangetoond in de groep vrouwen met een MCV van 90 of hoger.

TABEL 5: Het verband tussen hemoglobinegehalte, Mean Corpuscular Volume en percentage klachten.

klacht	Hb < 7.0 mmol/l		Hb > 7.0 mmol/l	
	MCV < 90 n = 56	MCV > 90 n = 73	MCV < 90 n = 112	MCV > 90 n = 210
moeheid	68 (%)	79 (%)*	57 (%)	69 (%)
hoofdpijn	14	11	13	14
duizeligheid	27	23	18	18
hartkloppingen	14	22 **	9	11
kortademigheid	41	49 *	33	36
bleekheid	18	32	26	26
concentratie- problemen	5	15	10	15
prikkelbaarheid	32	44	47	44
slecht slapen	11	15	11	17
neerslachtigheid	15	23	24	19
kuitkrampen	50	60	52	44

\* :  $p < 0.05$

\*\* :  $p < 0.01$

(Chikwadraat toets, tweezijdig, voor de vergelijking met de proportie vrouwen met klachten onder alle overigen).

## Discussie.

De conclusie van het onderzoek is, dat er bij zwangeren sprake lijkt te zijn van een verband tussen het voorkomen van bepaalde klachten en lage Hb- en Ht-waarden (reeds vanaf 7.0 mmol/l, respectievelijk 32%). In ons onderzoek kwamen vooral de volgende klachten naar voren: moeheidsklachten (erger dan vóór de zwangerschap), hartkloppingen, duizeligheid, kuitkrampen vooral als deze in de ochtend optreden. Na opsplitsing van de vrouwen in groepen met hogere respectievelijk lagere MCV- waarden (afkappunt 90 fl), werden verbanden tussen klachten en lage Hb- en Ht- concentraties alleen aangetoond in de eerste groep. Dit is wellicht een aanwijzing, dat foliumzuurtekort een rol speelt bij het ontstaan van klachten. Sommige klachten lijken (ook) met een hoger Hb samen te hangen, zoals hoofdpijn en onverklaarde prikkelbaarheid. Indien als verklaring voor prikkelbaarheid als psychosociaal geëvalueerde problemen worden aangegeven, dan is er juist meer kans op een laag Hb. De vraag is dan gewettigd, in hoeverre deze prikkelbaarheid werkelijk psychosociaal van aard is.

In tegenstelling tot wat wij verwachtten hangt moeheid die toeneemt in de loop van de dag, niet samen met een laag Hb. Waarschijnlijk is de invloed van de extra belasting in de loop van de dag op zichzelf voldoende verklaring voor deze klachten.

Het feit dat kuitkrampen in de ochtend zo duidelijk verband houden met lage hematologische waarden, werpt een nieuw licht op de genese van deze klachten. Wellicht betreft het hier een ander type klachten, dan de krampen die alléén 's nachts, in horizontale ligging optreden. Het kan zijn dat, als de zwangere 's morgens 'op gang' komt het hemoglobinegehalte net te kort schiet en dat een kortdurende relatieve ischemie ontstaat. Ook het feit, dat na lang in horizontale houding liggen een extra hemodilutie optreedt (5), kan hieraan bijdragen.

Klachten, die in meer of mindere mate gerelateerd kunnen worden aan lage Hb-waarden, zijn overigens 'gewoon' tijdens de zwangerschap. Sommige klachten (moeheid, kuitkrampen) komen bij de helft of meer van de vrouwen voor, en slechts zeer weinigen hebben géén klachten.

Indien ijzertabletten niet of niet alleen routinematig (c.q. at random), maar ook naar aanleiding van klachten zijn gebruikt, is te verwachten, dat de door ons gevonden verbanden sterker zouden zijn

geweest indien voor de 30e week in het geheel niet gesuppleerd was. Immers indien vooral vrouwen die tevoren reeds klachten hadden staal-tabletten -en daarmee wellicht een hoger Hb- hadden gekregen, dan zou het daarna ontdekken van een verband tussen een laag Hb en klachten moeilijker zijn. Niettemin vonden wij duidelijke relaties, met en zonder controle voor het gebruikt hebben van staaltabletten.

Is dit alles (toch weer) een pleidooi voor, eventueel routinematige, ijzersuppletie? Op grond van het hier beschreven onderzoek is die vraag niet te beantwoorden. Weliswaar staat vast, dat ijzersuppletie - overigens ook nog bij hoge Hb-waarden - de erythropoiese bevordert (6). Of hiermee ook een gunstig effect op subjectieve klachten bereikt wordt, is voor zover wij weten nog niet onderzocht. Daarnaast is het denkbaar dat met het oog op het verloop van de zwangerschap een matig hoog Hb-gehalte ongunstiger is, dan een matig laag Hb (7-10). Uit ons onderzoek zou, louter vanuit een oogpunt van subjectieve klachten, een ondergrens van 7.0 mmol/l voor het Hb, na de fysiologische hemodilutie, kunnen worden afgeleid. Deze is iets hoger dan de 6.3 à 6.5 mmol/l die is aanbevolen vanuit een oogpunt van een optimale zwangerschapsuitkomst (7,10).

Bespreking behoeft nog de vraag of de hier beschreven bevindingen te rijmen zijn met de resultaten van ons onderzoek bij niet-zwangeren in de huisartspraktijk, waarbij géén verbanden tussen klachten en het Hb werden gevonden(1). Men zou kunnen wijzen op het verschil in de gehanteerde methoden. Bij het huisartsenonderzoek werden de patienten door de huisarts ondervraagd, terwijl de zwangeren zelf het vragenformulier invulden. Daar in beide gevallen het Hb niet bekend was op het moment van de registratie van de klachten is systematische vertekening in beide gevallen echter onwaarschijnlijk. Een aannemelijke verklaring is, dat wij in de huisartspraktijk onder de moeheidspatienten geen Hb-waarden onder de 7.0 mmol/l vonden. Het lijkt er nu op dat, althans bij zwangeren, een Hb lager dan 7.0 mmol/l wel met klachten samenhangt. Hemodilutie kan dan weliswaar fysiologisch zijn, er treden wel wat meer klachten bij op.

Wij keren tenslotte terug naar de vraag in hoeverre de bepaling van de hematologische parameters gedurende de zwangerschap op geleide van subjectieve klachten kan geschieden. Wij zouden hierop voorlopig willen antwoorden: als men het belangrijk vindt een laag Hb bij zwan-

geren op te sporen, kan men de bepaling niet alleen op geleide van klachten uitvoeren. De redenen zijn, dat klachten ook bij zwangeren met een hooë Hb vrij vaak voorkomen, dat veel vrouwen met een laag Hb toch bepaalde klachten niet hebben, en dat het onderscheidend vermogen van geen van de klachten erg groot is. Omdat, zoals beschreven, bepaalde klachten wel verband houden met een laag Hb of Ht, kunnen deze wel als indicatie gelden voor een extra bepaling, los van de routinematige controles. Het verstrekken van staaltabletten louter naar aanleiding van klachten, zonder hematologische controle vooraf, lijkt in elk geval ongerechtvaardigd gezien de matige voorspellende waarde van klachten.

#### LITERATUUR.

1. Knottnerus JA, Knipschild PG, Van Wersch J, Sijstermanns AHJ. Onverklaarde moeheid en hemoglobinegehalte, een onderzoek vanuit de huisartspraktijk. Ned Tijdschr Geneesk 1986; 130: 402-405.
2. Wintrobe MM (ed.). Clinical Haematology. New York 1981, pp. 549-551.
3. Brockerhoff P, Kurtenbach MI, Stark W, Schiketanz KH, Freidberg V, Rathgen GH. Klinisch-chemische und hematologische Parameter im Schwangerschaftsverlauf. Gynäk Rdsch 1981; 21: 101-123.
4. BMDP Statistical Software, University California Press, Berkely 1983.
5. Kreutzer HH, Pennings AW. Het schommelende Hb-gehalte. Ned Tijdschr Geneesk 1976; 120: 144-146.
6. Taylor DJ, Lind T. Red cell mass during and after pregnancy. British Journal of Obstetrics and Gynaecology 1979; 86: 364-370.
7. Garn SM, Ridella SA, Petzold AS, Falkner F. Maternal hematologic levels and pregnancy outcomes. Seminars in Perinatology 1981; 5: 155-162.
8. Huisjes H. Het hemoglobine-gehalte in de zwangerschap. Ned Tijdschr Geneesk 1985; 129: 1323-1324.
9. Huisman A, Aarnoudse JG. Increased 2nd trimester hemoglobin concentration in pregnancies later complicated by hypertension and growth retardation. Acta Obstetrica et Gynaecologica Scandinavica. In press.
10. Murphy JF, O'Riordan J, Newcombe RG, Coles EC, Pearson JF. Relation of haemoglobin levels in first and second trimester to outcome of pregnancy. The Lancet 1986; 1: 992-994.



Appendix.

Stel:  $P(D^+)$  = de kans op ziekte D in een bepaalde (onderzoeks) populatie, zonder dat verdere diagnostische gegevens bekend zijn (prior kans).

$P(D^+|X)$  = de kans op ziekte D in deze populatie, gegeven het bekend zijn van testuitslag X (posterior kans).

In de logistische regressiefunctie geldt voor het verband tussen  $P(D^+|X)$  en X.

$$\text{Ln } \frac{P(D^+|X)}{1 - P(D^+|X)} = \text{natuurlijke logaritme van de posterior odds} = A + B X$$

waarbij A het intercept is en B de logistische regressiecoëfficiënt van X.

Om in het intercept de prior kans te verwerken, dient men als nieuw intercept te stellen:

$$A^1 = \text{Ln } \frac{P(D^+)}{1 - P(D^+)} = \text{natuurlijke logaritme van de prior odds.}$$

De functie zal na de conversie van het intercept dezelfde blijven indien X wordt vervangen door

$$X^1 = X - \frac{\text{Ln}(\text{prior odds}) - A}{B} \quad (2)$$

Het is nu gemakkelijk in te zien dat indien:

$$X = X_0 = \frac{\text{Ln}(\text{prior odds}) - A}{B} \quad (3)$$

geldt: posterior odds = prior odds

Dat wil zeggen: het aannemelijkheidsquotiënt van de uitslag met waarde  $X_0$  is gelijk aan 1.0.

Voor dichotome testvariabelen kan  $X_0$  een virtuele waarde van X zijn tussen 0 en 1.

De regressievergelijking wordt dan:

$$\ln (\text{posterior odds}) = \ln (\text{prior odds}) + B (X - X_0) \quad (4)$$

Dat wil zeggen:

$$\text{posterior odds} = \text{prior odds} * e^{B (X - X_0)} \quad (5)$$

Het aannemelijkheidsquotiënt (likelihoodratio, LR) van de testuitslag is dan:

voor de positieve uitslag ( $X = 1$ ):

$$LR+ = e^{B(1-X_0)} = e^{B-(A-\ln(\text{prior odds}))} \quad (6)$$

en voor de negatieve testuitslag ( $X = 0$ )

$$LR- = e^{-BX_0} = e^{A-\ln(\text{prior odds})} \quad (7)$$

$$\text{Tevens geldt } \frac{LR+}{LR-} = e^B, \text{ dat is: de Odds Ratio} \quad (8)$$

Ofwel:  $\frac{\text{posterior odds bij testuitslag } X = 1}{\text{posterior odds bij testuitslag } X = 0}$

De variabele  $X$  kan behalve een dichotome test ook een test met een uitslag op een continue schaal representeren. In (4) en (5) gaat het dan niet om het aannemelijkheidsquotiënt van een positieve of negatieve testuitslag, maar om het aannemelijkheidsquotiënt voor de (exakte) waarde  $X$ .

Door de ad (2) genoemde conversie van de onafhankelijke variabele  $X$  te verrichten kan men dus met behulp van logistische regressie-analyse rechtstreeks aannemelijkheidsquotiënten bepalen, en wordt het intercept equivalent aan de natuurlijke logaritme van de prior odds.

Door tijdens de data-analyse eerst voor iedere testvariabele  $X_i$  in een univariate regressie-analyse  $X_{i0}$  te bepalen, en deze vervolgens in de

multivariabele analyse van  $X_i$  af te trekken via de transformatie  $X_i = X_i - X_{i0}$ , krijgt men de volgende vergelijking.

$$\ln(\text{posterior odds}) = A^1 + B_1(X_1 - X_{10}) + B_2(X_2 - X_{20}) + \dots + B_i(X_i - X_{i0}) \quad (9)$$

Waarbij  $e^{A^1} = \text{prior odds}$

$$\text{en } e^{B_i(X_i - X_{i0})} = \text{aannemelijkheidsquotiënt voor variabele } X_i \quad (10)$$

met waarde  $X_i$   
gecorrigeerd voor de invloed van andere onafhankelijke variabelen.

Aangenomen wordt dan, dat de  $X_0$  waarden voor de diverse variabelen onafhankelijk van elkaar zijn. Deze aanname kan gecontroleerd worden via het onveranderd blijven van het intercept bij toevoeging van een nieuwe variabele.

Via dit model kan men, aldus, beschikbare diagnostische informatie gebruiken in verschillende combinaties, om uitgaande van de a priori kans de a posteriori kans te schatten, waarbij de betreffende aannemelijkheidsquotiënten gecorrigeerd zijn voor andere onafhankelijke (test) variabelen.

In feite re presenteert het model een 'probability tree' waarbij men, uitgaande van de prior kans, via toevoeging van een bepaalde set diagnostische informatie uiteindelijk een bepaalde "path probability" bereikt.

---

DIAGNOSTISCHE CONCLUSIES VAN DE HUISARTS NAAR AANLEIDING VAN ONVERKLAARDE MOEHEID.

J.A. Knottnerus - R. Starmans - A. Vissers

Samenvatting.

In een na-onderzoek, 1 jaar na het eerste contact vanwege moeheidsklachten die in eerste instantie onverklaard bleven, werden via het kaartsysteem en de huisarts zelf geïnventariseerd: de diagnostische conclusies van de huisarts naar aanleiding van de moeheidsepisode (n=161), alsmede alle diagnostische conclusies die binnen 1 jaar naar voren kwamen (N = 166). Naar aanleiding van de moeheidsepisode werd in 17% van de gevallen een somatische diagnose gesteld en bij 16% kwam een psychosociaal probleem naar voren. Op hogere leeftijd werden vaker somatische diagnosen gesteld. Bij 70% van de patiënten werd géén diagnostische conclusie geformuleerd naar aanleiding van de moeheidsepisode. Bij 55% van de patiënten was ook na verloop van een jaar geen diagnose of probleem vastgesteld.

Capaciteitsgroep Epidemiologie en Capaciteitsgroep Huisartsgeneeskunde, Rijksuniversiteit Limburg, Maastricht.

## Inleiding.

Als een patiënt met de klacht moeheid de huisarts bezoekt, zal dikwijls al tijdens het eerste consult een voor de hand liggende verklaring voor de klacht worden gevonden. Er blijft echter een groep over, bij wie aan het eind van het consult geen bevredigende conclusie is getrokken. Bij deze groep patiënten met 'onverklaarde moeheid' wordt vervolgens vaak aanvullende diagnostiek verricht (1-3). In een eerder onderzoek werd geen verband gevonden tussen onverklaarde moeheid en de waarde van hematologische parameters. Bovendien bleek dat de voorspellende waarde van onverklaarde moeheid ten aanzien van het voorkomen van anemie zeer laag was (4). Op grond van deze resultaten zou men het belang van onverklaarde moeheid als "diagnosticum" kunnen relativeren. Er ontbreekt echter nog een schakel, namelijk die van de - uiteindelijk - naar aanleiding van het eerste consult vastgestelde diagnosen of problemen.

Over de diagnostische betekenis van de klacht moeheid zijn tot nu twee huisartsgeneeskundige studies gepubliceerd. Beide waren retrospectief: de (ingangs)klacht werd achteraf ontleend aan patiëntenkaarten. Het eerste onderzoek betrof 176 patiënten met een "isolated diagnosis of fatigue" (5), en het tweede 57 patiënten met "fatigue of uncertain etiology", waarbij de oorzaak van de klacht onzeker was na anamnese en fysisch diagnostisch onderzoek (6). Men kwam in deze studies uiteindelijk in 40 à 50% van de gevallen tot somatische diagnosen en eveneens in 40% à 50% tot psychische problemen, al dan niet in combinatie. In 10 à 20% van de gevallen werd ook na follow-up geen definitieve diagnose gesteld. Gezien het retrospectieve karakter van beide studies moeten selectieve uitval en registratie van patiënten niet uitgesloten worden geacht. Hierdoor kan de hoeveelheid achterliggende pathologie zijn beïnvloed. Tevens werd de onverklaarde moeheid pas - soms geruime tijd - achteraf als zodanig gedefinieerd. Bovendien betrof het Amerikaanse onderzoeken, en het is bekend dat de Amerikaanse Family Practice niet vergelijkbaar is met de Nederlandse huisartspraktijk. Wij stelden een na-onderzoek in bij de moeheidspatiënten die in het eerste half jaar van 1984 waren opgenomen in het onderzoek "onverklaarde moeheid en het hemoglobinegehalte" (4). De vraagstelling was

hoe vaak de huisarts naar aanleiding van in eerste instantie onverklaarde moeheidsklachten uiteindelijk tot diagnostische conclusies komt en wat de aard is van deze conclusies.

#### Methode.

Met 9 huisartsen werd afgesproken dat gedurende een half jaar alle patiënten in het onderzoek zouden worden opgenomen,

- o die moeheidsklachten presenteerden waarvoor de huisarts na anamnese en fysische diagnostiek, verricht tijdens het eerste consult, geen verklaring kon vinden;
- o die tevoren niet bekend waren met pathologie die in verband kon staan met moeheid of het hemoglobinegehalte;
- o die in de 12 maanden tevoren geen moeheidsklachten bij de huisarts hadden gepresenteerd;
- o bij wie de huisarts een hemoglobinebepaling geïndiceerd achtte.

Deze toelatingscriteria werden van te voren mondeling doorgenomen en schriftelijk aan de huisartsen ter hand gesteld. Gedurende het onderzoek was er minstens eenmaal in de twee weken contact met de huisartsen over de voortgang van het onderzoek.

Aan de in het onderzoek opgenomen patiënten werd tijdens het eerste contact door de huisarts onder meer gevraagd naar de ernst (gemeten op een driepuntsschaal) en de duur van de klacht.

Van deze patiënten werd op de volgende wijze follow-up informatie verzameld.

- o Om vast te stellen welke diagnostische conclusie naar aanleiding van de moeheidsepisode getrokken werd, bestudeerden de auteurs eerst van iedere patiënt de patiëntenkaart van de huisarts en eventuele specialistische gegevens, één jaar na het eerste consult wegens moeheid. Op een standaardformulier werden de uiteindelijke diagnostische conclusies van de huisarts genoteerd. Er werden hierbij geen expliciete diagnostische criteria geformuleerd. Aangesloten werd bij de conclusie van de huisarts. In de regel waren deze direct aan de kaart te ontlezen. De diagnose "ijzergebrek" werd nooit expliciet genoemd. Daarom werd tot de diagnostische conclusie "ijzertherapie" besloten als de kaart alleen vermeldde dat ijzermedicatie werd verstrekt.

o Alle patiënten werden vervolgens met de eigen huisarts besproken. Daarbij werden de gegevens van de patiëntenkaart zonodig aangevuld. Voorts werd aan de huisarts gevraagd of zich bij de patiënten in het jaar na het eerste contact belangrijke nieuwe gezondheidsproblemen hadden voorgedaan, ook los van de moeheidsklachten. Ook over een aantal patiënten van wie geen patiëntenkaart was aangetroffen, kon de huisarts deze informatie verstrekken. Omdat de relatie tussen eventuele diagnostische conclusies en de moeheidsepisode bij deze (5) patiënten niet via de kaart te controleren was, werden zij van dat deel van het onderzoek uitgesloten.

### Resultaten.

In de onderzoeksperiode werden 174 patiënten in het onderzoek opgenomen. Van 161 van deze patiënten was het mogelijk om via het kaartsysteem (en de informatie van de huisarts zelf) alle follow-up gegevens te verzamelen. Van 5 andere (vrouwelijke) patiënten was de patiëntenkaart niet te vinden, maar van hun toestand na één jaar was de huisarts wel op de hoogte. Van 4 patiënten was bekend dat ze verhuisd waren, en volgens de huisarts had zich bij hen tot het moment van verhuizing geen relevante gezondheidsproblematiek voorgedaan. Over 4 patiënten kon noch via de kaart en noch via de huisarts informatie verkregen worden. Onder de 166 (161+5) patiënten van wie follow-up gegevens werden verkregen waren 39 mannen (gemiddelde leeftijd 41 jaar, range 18-82 jaar) en 127 vrouwen (gemiddelde leeftijd 38 jaar, range 18-79 jaar).

Gemiddeld hadden de patiënten 5.1 contacten met hun huisarts in het follow-up jaar (5.5 bij vrouwen en 3.7 bij mannen), waarvan gemiddeld 1.3 aan de moeheidsepisode werd besteed. Bij 13% van de opgevolgde patiënten was na een jaar de moeheidsklacht nog steeds aan de orde in het contact met de huisarts.

### Diagnostische conclusies naar aanleiding van de moeheidsklachten.

In tabel 1 zijn de conclusies naar aanleiding van de moeheidsepisode vermeld voor de 161 patiënten van wie de patiëntenkaart beschikbaar was, onderverdeeld naar geslacht en leeftijd (jonger dan 50 jaar, en 50 jaar en ouder). De conclusies zijn gegroepeerd in somatische diagnosen en psychosociale problemen (indien er een toestandsbeeld of een concreet probleem werd genoemd).

Tabel 1: Diagnostische conclusies van de huisarts naar aanleiding van in eerste instantie onverklaarde moeheidsklachten.

	mannen			vrouwen		
	< 50 jr n = 25	> 50 jr n = 13	totaal n = 39	< 50 jr n = 92	> 50 jr n = 30	totaal n = 122
<b>SOMATISCH</b> (aantal pat.)	(n=4)	(n=4)	(n=8)	(n=11)	(n=8)	(n=19)
bovenste lucht- weginfectie	2		2	3		3
hart/vaatpathol.						
angina pectoris		1	1		1	1
hypertensie				1		1
hypotensie					1	1
decomp. cordis					1	1
ijzer therapie				1	3	4
maag/darmpathol.						
maagulcus	1	1	2			
gastritis		1	1			
genitaal infect		1	1	2		2
klachten bewe- gingsapparaat				1	1	2
artrose				1		2
M Pfeiffer				1		1
CARA					1	1
allergie				1		1
urineweginfectie	1		1			
<b>totaal diagnosen</b>	<b>4</b>	<b>4</b>	<b>8</b>	<b>11</b>	<b>8</b>	<b>19</b>
<b>PSYCHOSOCIAAL</b> (aantal pat.)	(n=3)	(n=5)	(n=8)	(n=13)	(n=4)	(n=17)
depressief	1	1	2	2	2	4
gezin/relatie	1	1	2	2	1	3
werk		3	3	1		1
overige	2	2	4	11	4	15
<b>totaal problemen</b>	<b>4</b>	<b>7</b>	<b>11</b>	<b>16</b>	<b>7</b>	<b>23</b>
<b>GEEN DIAGNOSTI- SCHE CONCLUSIE</b> (aantal pat.)	(n=19)	(n=5)	(n=24)	(n=69)	(n=19)	(n=88)



Uit de tabel blijkt dat bij 17% (nl (8+19)/161) van de patiënten somatische diagnoses gesteld werden. Bij 16% (nl (8+17)/161) kwam een psychosociaal probleem naar voren. Er werden bij mannen relatief wat vaker somatische diagnoses en psychosociale problemen vastgesteld, maar deze verschillen zijn niet significant (chi-kwadraat toets). De leeftijd is van invloed, zoals ook blijkt uit tabel 2. Bij patiënten van 50 jaar en ouder wordt over het geheel genomen ruim tweemaal zo vaak een somatische diagnose gesteld als bij jongere patiënten ( $p < 0.05$ , chi-kwadraat toets). Deze tendens doet zich bij psychosociale problemen alleen bij mannen voor.

TABEL 2: Percentage moeheidspatiënten bij wie naar aanleiding van de moeheidsepisode een bepaald probleem wordt vastgesteld, naar leeftijd en geslacht.

Aard probleem	mannen		vrouwen		mannen en vrouwen	
	< 50 jr n = 26	> 50 jr n = 13	< 50 jr n = 92	> 50 jr n = 30	< 50 jr n = 118	> 50 jr n = 43
Somatisch	15%	31%	12%	27%	13%	28%
Psycho-sociaal	12%	38%	14%	13%	14%	21%

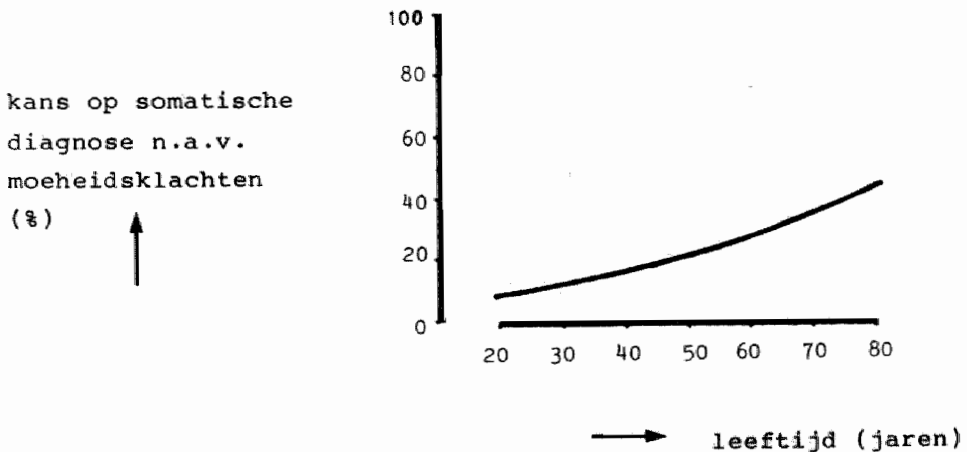
Gezondheidsproblemen kunnen in combinatie voorkomen. De grootste groep patiënten bij wie een probleem naar voren komt betreft een geïsoleerde somatische diagnose (15%). Bij 11% komt alleen een psychosociaal probleem voor. Combinaties worden slechts in 5% van de gevallen gevonden, terwijl bij 70% van de moeheidspatiënten (112/161) géén diagnostische conclusie naar voren komt (tabel 1).

Wij vonden geen relevant verband tussen de duur van de klachten vóór het eerste contact en de kans op het vinden van een somatische aandoening of een psychosociaal probleem. Ook tussen de gerapporteerde ernst van de moeheid (tamelijk moe, moe, erg moe) en de kans op gezondheidsproblemen vonden wij geen samenhang.

Tot nu toe stonden wij nog niet stil bij de aard van de diagnostische conclusies. Tabel 1 geeft aan dat deze zeer gevarieerd is. Als men de somatische diagnoses hergroepeert in infectieuze en niet-infectieuze aandoeningen, valt op dat de eerste vooral bij jongeren (< 50 jaar)

voorkomen (60% van de bij hen gestelde somatische diagnosen). Bij de ouderen overheersen meer chronische aandoeningen (92% van de somatische diagnosen). Bij slechts 4 patiënten (allen vrouw) werd ijzer voorgeschreven. Hierbij kan bedacht worden dat geen van hen bij het eerste consult een Hb lager dan 7.5 mmol/l had (4).

Naar aanleiding van de ziekte-episode die begon met de presentatie van moeheid werden 20 patiënten (12%) verwezen naar een medisch specialist. De longarts, de internist en de cardioloog namen hiervan driekwart voor hun rekening. Zoals te verwachten wordt in deze groep relatief vaak een somatische diagnose gesteld (bij 12 patiënten, overeenkomend met 60% van de verwezen gevallen, versus 17% van de totale groep). Dit demonstreert dat ten gevolge van selectie specialisten bij patiënten met moeheidsklachten veel vaker somatische diagnosen aantreffen dan huisartsen.



Figuur 1: Verband tussen de leeftijd en de kans dat er een somatische diagnose gesteld wordt in aansluiting op onverklaarde moeheid.

---

logistische functie:		
0 intercept	-3,04	SE: 0,61
0 logistische regressie		
coëfficiënt voor leeftijd	0,035	SE: 0,013

---

Door middel van een (forward) stepwise multiple logistische regressie-analyse werd onderzocht in hoeverre diverse variabelen de kans beïnvloeden dat er een somatische diagnose gesteld wordt. Afhankelijke variabele was de (logit van de) kans op een somatische diagnose. Onafhankelijke variabelen waren het geslacht, de leeftijd, de duur van de moeheidsklachten en de gerapporteerde ernst ervan. Behalve de leeftijd bleek in deze analyse geen van de andere variabelen van invloed. In figuur 1 is, op grond van deze analyse, de kans op een somatische diagnose als functie van de leeftijd weergegeven.

Situatie na één jaar follow-up.

In de loop van het jaar na het eerste contact kwamen diverse gezondheidsproblemen naar voren los van de moeheidsepisode. Van de 5 vrouwelijke patienten van wie geen patientenkaart gevonden was (zodat het precieze verband met de moeheidsepisode niet kon worden nagegaan) was bij de huisarts wel follow-up informatie over geconstateerde problemen bekend. Onder de 166 (161+5) opgevolgde patiënten bleek de groep met een op zich zelf staande somatische diagnose nog steeds de grootste. In totaal werd uiteindelijk bij 28% van de patiënten een somatische diagnose gesteld en kwam bij 25% een psychosociaal probleem naar voren. Bij 55% (92/166) van de opgevolgde patiënten is ook na verloop van een jaar noch een somatische diagnose noch een psychosociaal probleem vastgesteld.

De meeste van de later gestelde diagnoses betreffen intercurrente 'minor ailments' zoals blaasontsteking en huidafwijkingen. Ernstige later gevonden pathologie, die niet in het kader van de moeheidsepisode naar voren was gekomen (maar achteraf mogelijk toch daarmee in verband zou zijn te brengen), wordt gemeld in tabel 3, onderverdeeld naar leeftijd en geslacht. Deze deed zich vooral voor bij oudere vrouwen. Een patiënte overleed, en bij haar werd bij obductie een gemetastaseerd maagcarcinoom gediagnosticeerd. Bij een jongere vrouw kwam een mammacarcinoom aan het licht. De invloed van de leeftijd op de kans dat er na 1 jaar een somatische diagnose aan het licht is gekomen, is minder duidelijk dan de overeenkomstige kans naar aanleiding van de moeheidsklachten, en statistisch niet significant.

TABEL 3: Ernstige pathologie die bij follow-up van een jaar werd gevonden, los van de initiële moeheidsklachten.

<u>mannen</u>		<u>vrouwen</u>	
< 50 jaar	≥ 50 jaar	< 50 jaar	≥ 50 jaar
geen	CVA bij diabetes	mammacarcinoom	reumatoïde artritis
		hypertensie	angina pectoris
		Ig A deficiëntie	nierinsufficiëntie
			CARA
			gemetastaseerd maagcarcinoom (obductie) †

### Discussie.

Evenals Morrison (5) en Katerndahl (6) vonden wij dat naar aanleiding van in eerste instantie (na anamnese en fysische diagnostiek) onverklaarde moeheidsklachten een grote variëteit aan somatische en niet-somatische problemen door de huisarts wordt vastgesteld. In onze patiëntengroep werd echter veel minder vaak zo'n probleem gevonden. Bij 70% van de patiënten kwam voor de moeheidsepisode geen verklaring naar voren, en bij 55% werd in de loop van een jaar follow-up geen diagnose of probleem vastgesteld. Somatische en niet-somatische problemen komen bij ongeveer evenveel personen voor, al dan niet in onderlinge combinaties. Het percentage patiënten bij wie ook in tweede instantie geen verklaring werd gevonden was in onze studie beduidend hoger dan bij Morrison (5) en Katerndahl (6). Dat kan liggen aan verschillen tussen de patiëntenpopulaties, de diagnostische criteria, of tussen de dokters, maar ook aan de onderzoeksmethode of de definitie van onverklaarde moeheid. In onze studie werd op grond van vooraf vastgestelde criteria bepaald welke moeheidspatiënten in het onderzoek werden opgenomen, en er was weinig uitval. In de andere onderzoeken bepaalde men achteraf op grond van informatie op de patiëntenkaarten bij wie sprake was geweest van onverklaarde moeheid. De volledigheid van follow-up is moeilijk te beoordelen. Onze cijfers

geven onzes inziens dan ook een betrouwbaarder beeld. Daarbij dient bedacht te worden dat onze studie patiënten betrof, bij wie in eerste instantie een verklaring ontbrak maar wel een hemoglobine bepaling geïndiceerd geacht werd.

Wat betreft de aard van de uiteindelijk gestelde somatische diagnoses vormen, evenals bij de genoemde Amerikaanse auteurs infectieziekten de grootste subgroep. Ook hart-vaataandoeningen leveren een substantiële bijdrage. IJzertherapie werd in onze patiëntengroep, waarvoor bij het eerste consult een hemoglobinebepaling geïndiceerd werd geacht, betrekkelijk zelden voorgeschreven. Er worden 2 maligniteiten gediagnosticeerd in de follow-up periode, waarvan één pas bij obductie.

De kans op het stellen van een somatische diagnose werd duidelijk beïnvloed door de leeftijd van de patiënt. Het geslacht speelde hierbij geen rol, evenals de gerapporteerde ernst van de klacht. Bij dit laatste gegeven is uiteraard een zekere subjectiviteit in het spel. Maar omdat op het moment van de beoordeling van de ernst de klacht nog onverklaard was, lijkt systematische vertekening ten aanzien van de kans op een somatische diagnose onwaarschijnlijk. Wat betreft de duur van de klacht konden wij niet de bevinding van anderen bevestigen, dat recent ontstane klachten vaker op een somatische diagnose wijzen en langer durende klachten meer in verband staan met psychische en psychosociale problemen. (5,6)

Het is natuurlijk mogelijk dat een aantal diagnoses en problemen niet aan de arts worden voorgelegd of door hem niet worden herkend. Daarbij hoort ook de kanttekening te worden geplaatst, dat de patiënten ná het eerste consult niet aan gestandaardiseerde follow-up diagnostiek werden onderworpen op grond van uniforme criteria. Dit zou gezien de grote variëteit van de mogelijke problemen ook nauwelijks realiseerbaar zijn geweest. Wij achten het evenwel aannemelijk, dat in elk geval de ernstige, medisch relevante problematiek na verloop van een jaar wel bekend is geworden aan de huisarts. De conclusie lijkt gewettigd dat achter moeheidsklachten, waarvoor de huisarts in eerste instantie door middel van anamnese, fysische diagnostiek én hematologisch onderzoek geen verklaring vindt, ook later in de meeste gevallen geen ernstige problematiek blijkt schuil te gaan.

LITERATUUR.

1. Berris B, Rachlis A. Investigation of fatigue. Canadian Family Physician 1977; 23: 465-466.
2. Avonts D, Denekens J, Hendriks K, Neels H. De patiënt die moe is. Nieuw Kompas voor de Huisarts 1983; III/2: 1-9.
3. Solberg L.I. Lassitude, a primary care evaluation. JAMA 1984; 251: 3272-3276.
4. Knottnerus JA, Knipschild PG, Van Wersch J, Sijstermanns A. Onverklaarde moeheid en het hemoglobinegehalte, een onderzoek in de huisartspraktijk. Nederlands Tijdschrift voor Geneeskunde 1986; 130: 402-405.
5. Morrison JD. Fatigue as a presenting complaint in family practice. Journal of Family Practice 1980; 10: 795-801.
6. Katerndahl DA. Fatigue of uncertain etiology. Family Medicine Review 1983; 4: 26-38.

# Klachten op de borst: omgaan met onzekerheden

J. A. KNOTTNERUS, E. EBBENS, Th. M. E. GOVAERT EN  
C. A. DE GEUS\*

In acht huisartspraktijken werden gedurende een half jaar alle patiënten die zich met klachten op de borst tot de huisarts wendden, geregistreerd; bovendien werd zes maanden na het eerste consult informatie verzameld over de verdere lotgevallen van elke patiënt. Aldus werden gegevens verzameld over ruim 300 personen. Hoewel de huisarts in een kwart van de gevallen rekening hield met de mogelijkheid van een cardiale oorzaak van de klacht, bleek hij slechts een minderheid van deze patiënten direct te verwijzen; ten tijde van de *follow up* was de helft van de verdachte groep nog steeds volledig onder zijn hoede. Op dat moment had hij nog in 8 procent van de gevallen twijfels over de juistheid van de gestelde diagnose.

## Inleiding

De patiënt die zich tot de huisarts wendt met klachten op de borst, is vaak bang voor een hartkwaal en wil snel weten waar hij aan toe is. De arts wil hierna graag tegemoet komen, maar beseft tevens dat niet alleen het miskennen van een cardiaal probleem, maar ook het ten onrechte veronderstellen van zo'n probleem ingrijpende consequenties kan hebben voor het welzijn van de patiënt. Om hem bij zijn besluitvorming houvast te bieden, is er een toenemende aandacht voor de ontwikkeling van diagnostische protocollen. Deze zijn echter meestal gebaseerd op klinisch specialistische ervaring en slechts een enkele maal op onderzoek in de huisartspraktijk.<sup>1,2</sup> Daarnaast hebben, ook in dit tijdschrift, de mogelijkheden en beperkingen van de electrocardiografie in de huisartspraktijk speciale aandacht gekregen.<sup>3,5</sup>

In Maastricht zijn enkele aspecten van de feitelijke huisartsgeneeskundige betrokkenheid bij klachten op de borst bestudeerd. In dit verslag staat de volgende vraagstelling centraal:

- Wat is het patroon van nieuwe klachten op de borst dat aan de huisarts wordt gepresenteerd; hoe vaak wordt gedacht

aan, respectievelijk getwijfeld over een cardiale oorzaak van de klacht; in welke gevallen wordt verwezen, en welk beleid wordt uitgestippeld ten aanzien van niet verwezen patiënten;

- hoe verhoudt zich de beoordeling van de huisarts bij het eerste consult tot het oordeel dat hij later velt op grond van het klinische beloop?

## Methoden

Aan het onderzoek werd deelgenomen door acht huisartsen in de regio Zuid-Limburg, verbonden aan het Basisproject Registratie en Verslaggeving. Hun gezamenlijke praktijkpopulatie omvat 17.000 personen.

Het onderzoek heeft betrekking op de patiënten die zich in de tweede helft van 1983 tot deze huisartsen wendden met klachten op de borst. Uitgesloten werden patiënten met klachten die de huid of de mammae betroffen of die een direct duidelijke traumatische oorzaak hadden. Eveneens uitgesloten werden patiënten die de voorgaande twaalf maanden al eerder wegens klachten op de borst bij de huisarts waren geweest, en patiënten die vanwege dergelijke klachten onder behandeling waren (gewest) van een specialist.

Alle andere patiënten met klachten op de borst werden door de huisarts tijdens of direct na het eerste contact geregistreerd op daartoe speciaal ontworpen

standaardformulieren. De volgende gegevens werden opgetekend:

- basisgegevens als leeftijd en geslacht;
- de ingangsklacht en de codering hiervan volgens de Reason For Encounter Classificatie (RFE-C), laatste versie;<sup>6</sup>
- aspecten van het handelen van de arts met betrekking tot diagnostiek, therapie, verwijzing en vervolgspraken;
- de waarschijnlijkheidsdiagnose;
- de mate van verdenking op een cardiale oorzaak van de klacht, aangegeven op een driepuntsschaal (ja, twijfel, neen).

Omdat het feitelijk handelen van de arts onderzocht zou worden, zijn geen afspraken gemaakt over de diagnostische en therapeutische strategie.

Een half jaar na het eerste consult werd door de huisartsen over de patiënten informatie verzameld vanuit drie bronnen:

- de rapportage van de specialist, in geval van verwijzing;
- de eigen verslaggeving betreffende het klinische beloop;
- speciaal georganiseerde, meestal telefonische, contacten met alle in leven zijnde patiënten voor zover bereikbaar; hierbij werd een korte anamnese afgenomen.

Op basis van deze informatie gaven de huisartsen vervolgens voor alle patiënten opnieuw antwoord op de vraag, of de destijds gepresenteerde klacht van cardiale oorzaak was geweest (ja, twijfel, neen).

## Resultaten

### Eerste consult

Aan de criteria voldeden 325 patiënten (158 vrouwen en 167 mannen), van wie 285 op het spreekuur en 40 thuis werden gezien. De gemiddelde leeftijd was 47 jaar (S.D. 16 jaar). Eén patiënt overleed vlak voor de komst van de arts. In 92 procent van de gevallen ging het om pijnklachten, en meestal was die pijn niet nader te specificeren binnen de RFE-C (tabel 1).

De „carrière” van de bestudeerde groep tijdens het eerste consult is geschetst in de *figuur*. Hierbij is uitgegaan van de ingangsklachten, samengevoegd in drie categorieën:

- klachten mogelijk samenhangend met de circulatie (alle K-codes);
- niet nader te specificeren pijn op de borst (A11-code);
- overige klachten.

In 9 procent van de gevallen vermoedde de huisarts een cardiale oorzaak, terwijl hij bij 18 procent twijfel koesterde.

\* De auteurs zijn verbonden aan de Capaciteitsgroepen Gezondheidszorgonderzoek/Epidemiologie en Huisartsgeneeskunde van de Rijksuniversiteit Limburg.

Uiteraard vonden bijna alle directe veranderingen plaats in deze twee groepen. Omgekeerd werden echter lang niet alle verdachte gevallen verwezen: van de 28 patiënten bij wie de huisarts een cardiale oorzaak vermoedde, werden er slechts twaalf verwezen.

Naarmate de patiënten ouder waren, werd vaker aan een cardiale oorzaak gedacht. De gemiddelde leeftijden van de verdachte groep, de twijfelgroep en de niet-verdachte groep waren respectievelijk 66 (S.D. 14), 55 (S.D. 13) en 43 (S.D. 17) jaar. In de verdachte groep kwamen tweemaal zoveel mannen voor als vrouwen (18 versus 10), terwijl in de twijfelgroep de vrouwen oververtegenwoordigd waren (24 versus 33). De helft van de deelnemende huisartsen heeft een eigen ECG-apparaat. De keren dat direct werd verwezen, is geen enkele maal vooraf een ECG gemaakt. Bij 16 procent van de patiënten is na het consult een ECG gemaakt of aangevraagd; in de twijfelgroep was dat percentage zelfs 47 procent.

In die gevallen waarin een cardiale oorsprong werd overwogen, maar niet werd verwezen, volstond de huisarts meestal met het geven van voorlichting en/of medicatie. De meeste patiënten werden op korte termijn terugbesteld, maar in twaalf gevallen (20 procent van deze groep) werd het initiatief aan de patiënt overgelaten. De helft van deze patiënten had angina pectoris of een ritmestoornis.

#### Follow-up algemeen

Van zes patiënten zijn geen follow-up gegevens beschikbaar. Bij slechts één van hen was bij het eerste contact gedacht aan een mogelijke cardiale oorzaak, bij de overigen was daaraan in het geheel niet gedacht. Over de resterende 318 patiënten wordt in het volgende gerapporteerd: met 297 van hen is na een half jaar persoonlijk contact geweest.

Als we - met enige overdrijving - de beoordeling bij de follow-up beschouwen als de standaard, dan zien we in tabel 2 het volgende:

- De voorspellende waarde van een duidelijke positieve verdenking bedraagt iets meer dan 50 procent.
- De voorspellende waarde van een afwezige verdenking bedraagt 97 procent.
- In een kwart van de gevallen waarin de huisarts aanvankelijk twijfelt, blijft die twijfel bestaan.

In totaal bestaat na een half jaar nog twijfel bij 8 procent van de gehele onderzoekspopulatie; in absolute zin is

deze groep vrijwel even groot als de groep patiënten bij wie de huisarts zeker is van een cardiale oorzaak.

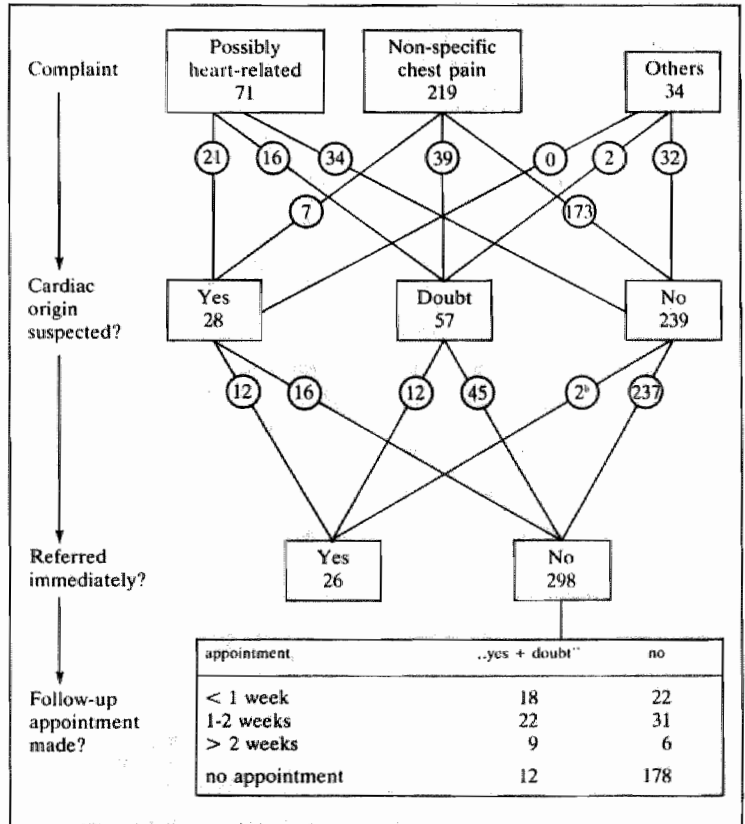
Van de 297 patiënten met wie bij de follow-up persoonlijk contact is ge-

weest, had 47 procent nog steeds last van de klachten. Daarbij maakte het niet veel uit, of men op dat moment tot de verdachte groep, de twijfelgroep of de niet-verdachte groep behoorde (per-

**Tabel 1 - Table 1.** Nieuw gepresenteerde klachten op de borst, geregistreerd volgens de RFE-classificatie (zes maanden, acht huisartspraktijken) - Newly presented chest complaints registered according to the RFE classification (six months, eight general practices).

Reason for encounter	Number (n = 325)	Percentage
A11 Pain chest, not elsewhere classified	219	67
K10 Pain attributed to heart and circulation	59	18
R10 Pain-respiration	20	6
K27 Fear of disease circulatory system	9	3
K11 Abnormal/Irregular heartbeat/pulse	4	1
A29 Other general symptoms/complaints	2	.5
P22 Psychophysiological problems	2	.5
Other symptoms/complaints (each registered once)	10	3

**Figuur 1 - Figure 1.** Wat gebeurt er met de patiënt die met klachten op de borst de huisarts consulteren in de loop van het eerste consult? - What happens with patients with chest complaints in the course of the first encounter with the general practitioner?



\* excluding the patient who died before the general practitioner arrived.

b patients with „pulmonary embolism“ and „cardiac neurosis“.



centages patiënten met klachten respectievelijk 56, 55 en 44 procent). Zes personen zijn binnen een half jaar overleden (tabel 3). De helft van hen behoorde tot de aanvankelijk niet-verdachte groep. In de meeste gevallen betwijfelde de huisarts bij de follow-up nog steeds de aanvankelijke klacht van cardiale aard was geweest.

#### Follow-up verwezen patiënten

Uiteindelijk zijn in de loop van het halfjaar follow-up nog 15 patiënten ver-

wezen, zodat het totaal aantal verwijzingen op 41 komt. Bij vergelijking van de diagnoses van de huisarts tijdens het eerste consult met de conclusies van de specialist bleek, dat van de zeven vermoede myocardiinfarcten er twee werden bevestigd; in de overige vijf gevallen concludeerde de specialist wel steeds tot ischémische hartziekte c.q. angina pectoris. Bij de 11 patiënten die direct waren verwezen vanwege angina pectoris, concludeerde de specialist slechts drie keer tot een cardiologische

afwijking. Daarnaast werd in de loop van het halve jaar via de specialist nog twee keer een infarct vastgesteld bij patiënten bij wie de huisarts aanvankelijk dacht aan angina pectoris respectievelijk hyperventilatie (tabel 4).

Uit tabel 4 blijkt verder dat er nog enkele malen een belangrijk verschil was tussen de eerste diagnose van de huisarts en de conclusie van de specialist. *Hiatus hernia* en *respiratoire infectie* traden op als voorlopers van de conclusie angina pectoris, terwijl achter een geval van *myalgie* een aneurysma dissekans bleek schuil te gaan. Daarbij moet worden aangetekend dat de eerste diagnose van de huisarts en de diagnose van de specialist in de later verwezen groep enige tijd uit elkaar liggen, en dat in die periode het beeld veranderd kan zijn. In een aantal gevallen liet de huisarts zich niet overtuigen door de diagnose van de specialist:

- Bij 12 patiënten concludeerde de spe-

**Tabel 2 - Table 2. Het aanvankelijke oordeel over de cardiale oorsprong van de klachten en het oordeel zes maanden later (percentages) - Initial judgement about the cardiac origin of the complaints and judgement six months later (percentages).**

First encounter	Follow-up		
	yes	doubt	no
— yes (n = 28)	54	14	32
— doubt (n = 56)	11	23	66
— no (n = 234)	.5	3	96.5

#### Synopsis

**Knottnerus JA, Ebbens E, Govaert ThME, De Geus CA. Chest problems: coping with uncertainties. Huisarts en Wetenschap 1985; 28: 159-64.**

**Introduction.** This study concerned some aspects of general practice performance in relation to chest problems. The problem definition was:

- What is the pattern of new chest problems presented to the general practitioner; how often is a cardiac origin of the problem suspected or doubted; what is the strategy planned with regard to unreferred patients?

- How does the general practitioner's assessment at the first encounter compare with his subsequent assessment on the basis of the clinical course?

**Methods.** Participants in the study were eight general practitioners with a total practice population of 17,000 persons. The study covered patients who presented in the latter half of 1983 with chest problems. Patients with problems involving the mammary skin or of evidently traumatic origin were excluded, as were patients who had presented with chest problems within the preceding twelve months and those who were being (had been) treated by a specialist in view of such problems. All other patients with chest problems were registered by the general practitioner on special forms during or immediately after the first encounter.

Six months after the first encounter the general practitioner again collected information on the patients involved. On the basis of this information the general prac-

titioners then answered the question, for all patients, whether the problem originally presented had been of cardiac origin.

**Results.** The criteria were met by 325 patients. Pain was reported in 92 percent of the cases, and in most cases this pain could not be further specified according to the Reason For Encounter Classification (RFE-C) (table 1). In 9 percent of cases the general practitioner suspected a cardiac cause, and in 18 percent he doubted this. Nearly all immediate referrals were of course made in this group. On the other hand, by no means all suspect cases were referred: of the 28 cases in which the general practitioner suspected a cardiac cause, only 12 were referred (figure).

Accepting the follow-up judgement as certain diagnosis, table 2 can be read as follows:

- the predictive value of a distinct positive suspicion amounted to slightly over 50 percent;
- the predictive value of an absent suspicion amounted to 97 percent;
- the general practitioner's initial doubt persisted in some 25 percent of the cases;
- doubt persisted after six months in a total of 8 percent of the study population; in absolute terms this group was virtually as large as the group in which the general practitioner was certain of a cardiac cause. In the end, 15 more patients were referred in the course of the six follow-up months, bringing the total number of referrals up to 41 (table 3). About 50 percent of the patients in whom a possible cardiac origin of the problem was suspected after these six months, were still entirely in the care of the general practitioner at that time.

Table 4 shows that, in the unreferred

group, an initial suspicion of ischaemic heart disease/angina pectoris was often subsequently doubted. A suspicion that the problems were caused by arrhythmias was usually rejected in retrospect. On the other hand, two general practitioners themselves subsequently diagnosed myocardial infarction on the basis of ECG findings.

**Discussion.** The predictive values listed in table 2 concur with the general findings of Wright and Kane, which indicate that in general practice the predictive value of a favourable prognosis always exceeds that of a less favourable prognosis.<sup>15</sup> Moreover, doubt proved to be a factor of importance, to which some predictive value might also be attached. Various other publications have shown that chest problems, even after detailed diagnostic work, are not always classifiable in the long run. A retrospective study in a number of general practices revealed that 50 percent of the cases of chest pain still had an unproven aetiology after six months<sup>16</sup>; and of a group of patients admitted to an intensive care unit with problems „suggestive of myocardial infarction“, 13 percent were discharged with the diagnosis *chest pain of obscure origin*.<sup>16</sup> The study clearly shows that general practitioners regard the guidance of patients with a possible cardiac disorder as part of their task, and accept a degree of uncertainty in these cases.

**Keywords:** Family practice; Diagnosis; Thoracic diseases.

**Address of authors:** Capaciteitsgroep Gezondheidszorgonderzoek / Epidemiologie Rijksuniversiteit Limburg, Postbus 616, 6200 MD Maastricht, The Netherlands.

cialist tot ischemische hartziekte (uitgezonderd infarct), terwijl de huisarts in drie gevallen bleef twijfelen.

• Bij 21 patiënten concludeerde de specialist dat er geen cardiologische afwijking was; ook nu bleef de huisarts in drie gevallen twijfelen.

• In totaal bleef bij acht patiënten (20 procent) twijfel bestaan.

#### Follow-up niet verwezen patiënten

Uit tabel 5 blijkt dat in de niet-verwezen groep een aanvankelijke verdenking op ischemische hartziekte/angina pectoris later vaak werd betwijfeld. Een vermoeden dat de klachten werden veroorzaakt door ritmestoornissen, is achteraf meestal verworpen.

Hier tegenover staat, dat twee huis-

artsen op basis van ECG-bevindingen later zelf een infarct vaststelden. Het beloop in deze gevallen was gunstig, en deze patiënten werden niet verwezen. Het totale aantal vastgestelde hartinfarcten in het vervolgde cohort kwam hiermee op zes, exclusief de patiënt die reeds was overleden voor het eerste contact.

**Tabel 3 - Table 3.** Sterfgevallen, opgetreden in het cohort waarvan follow-up informatie is verzameld (n = 318) - Deaths occurring in the cohort on which follow-up information has been collected (n = 318).

Patient	Information about first encounter			Follow-up information		
sex age	cardiac origin of complaint according to GP	tentative diagnosis	referred immediately?	referred?	cause of death	final judgement GP about cardiac origin of complaint
1. M 49	no	influenza	no	no	probably subarachnoidal haemorrhage after 5½ months	no
2. M 58	doubt	thoracic spondylosis	no	no	respiratory insufficiency after 6 months	doubt
3. M 67	no	virus infection	no	yes, for IHD/AP	acute death (probably infarction) after 6 months	doubt
4. M 58	yes	ventricular extrasystoles	no (pat. had already seen lung spec.)	no (already seen lung specialist)	pulmonary carcinoma with metastases after 5 weeks	doubt
5. M 81	yes	myocardial infarction	yes	yes	cardiac decompensation after 3 weeks	yes
6. F 83	no	myalgia	no	yes, for dissecting aneurysm	dissecting aneurysm after 5 months	doubt

**Tabel 4 - Table 4.** Waarschijnlijkheidsdiagnosen van de huisarts bij het eerste consult en de diagnosen van de specialist (n = 41) verwezen patiënten) - The general practitioner's tentative diagnoses at the first encounter and the specialist's diagnoses (n = 41 referred patients)

General practitioner's diagnosis	Specialist's diagnosis						total
	infarction	IHD/AP	arrhythmia	other cardiovasc. disease	no cardiolog. disease	no conclusion	
<b>Referred immediately</b>							
— infarction	2	5	—	—	—	—	7
— IHD/AP <sup>a</sup>	—	3 <sup>b</sup>	—	—	7	1	11
— other	—	—	1 <sup>c</sup>	1 <sup>d</sup>	5 <sup>e</sup>	1 <sup>f</sup>	8
<b>Referred later</b>							
— infarction	—	—	—	—	—	—	—
— IHD/AP	1	2	—	—	—	—	3
— other	1 <sup>g</sup>	2 <sup>h</sup>	—	2 <sup>i</sup>	7 <sup>a</sup>	—	12
<b>Total</b>	<b>4</b>	<b>12</b>	<b>1</b>	<b>3</b>	<b>19</b>	<b>2</b>	<b>41</b>

<sup>a</sup> ischaemic heart disease/angina pectoris (excluding infarction).

<sup>b</sup> including two patients referred in view of „imminent infarction“.

<sup>c</sup> general practitioner: pulmonary embolism; specialist: arrhythmia, possibly pulmonary embolism.

<sup>d</sup> both general practitioner and specialist: pericarditis.

<sup>e</sup> tentative diagnosis by general practitioner: chest

pain of obscure origin (twice), myalgia (twice), hyperventilation syndrome (once).

<sup>f</sup> tentative diagnosis by general practitioner: hyperventilation/cardiac neurosis.

<sup>g</sup> general practitioner at first encounter: hyperventilation.

<sup>h</sup> general practitioner's diagnoses at first encounter: 1) hiatus hernia, 2) respiratory infection.

<sup>i</sup> 1) general practitioner at first encounter: myalgia; specialist: dissecting aneurysm.

2) general practitioner at first encounter: gastric ulcer; specialist: peripheral circulatory disorder.

<sup>a</sup> the general practitioner usually did not venture beyond „non-specific chest complaints“; specialist's findings respectively: gastric carcinoma with liver metastases; bronchial asthma; costoclavicular syndrome; anaemia; no cardiovascular disease (not further qualified) (twice).

**Tabel 5 - Table 5.** Waarschijnlijkheidsdiagnosen van de huisarts bij het eerste consult en zijn oordeel na een half jaar over de cardiale oorzaak van de klacht ( $n = 277$  uiteindelijk niet verwezen patiënten) - General practitioner's diagnoses at first encounter and his judgement about the cardiac origin of the complaint six months later ( $n = 277$  patients ultimately not referred).

Diagnosis at first encounter	Judgement six months later			
	yes	doubt	no	total
— infarction	—	—	—	—
— IHD/AP	7	6	3	16
— arrhythmia	1	1	4	6
— chest complaint e.c.i.	—	3	40	43
— other	—	6	206	212
Total	8	16	253	277

In veruit de meeste gevallen waarin de huisarts aanvankelijk een klachtdiagnose of een niet-cardiologische diagnose had gesteld, verwierp hij uiteindelijk de mogelijkheid van een cardiologische oorzaak.

Van de patiënten met een angina pectoris of een bij de follow-up veronderstelde c.q. overwogen cardiale oorzaak, was een half jaar later ongeveer de helft geheel onder de hoede van de huisarts gebleven.

### Beschouwing

Het is om te beginnen nodig enige kanttekeningen te maken bij het onderzoek. Niet uitgesloten is dat het registreren het handelen van de huisarts heeft beïnvloed. Voorts konden het noteren van de klacht en van de waarschijnlijkheidsdiagnose niet onafhankelijk van elkaar plaatsvinden, en hetzelfde gold voor de beoordeling tijdens het eerste consult en die na een half jaar. Anderzijds gold hierbij het voordeel dat aangesloten werd bij de normale gang van zaken, waarin het oordeel van dezelfde huisarts over dezelfde patiënten een ontwikkeling doormaakt. Dit oordeel is bepalend voor zijn beleid. Evaluatie van verschillen tussen individuele huisartsen viel buiten het bestek van dit onderzoek. Wat betreft de incidentie van pijn op de borst komt ons resultaat redelijk overeen met dat van *Blalock*: 35 per 1000 personen uit de basispopulatie per jaar, versus 49 per 1000 per jaar. Daarbij moet worden bedacht dat de laatstgenoemde studie ook patiënten met duidelijke traumata en met reeds bekende klachten omvatte.<sup>7</sup>

De incidentie van angina pectoris in onze studie stemde goed overeen met observaties van anderen: men kan ervan uitgaan dat deze 3 tot 5 nieuwe gevallen per 1000 personen per jaar bedraagt.<sup>8,9</sup> Wat betreft het myocardiinfarct komen

wij lager uit: 1 versus 2 à 3 per 1000 per jaar. In onze studie waren echter patiënten die al bekend waren bij de cardioloog en/of reeds klachten op de borst hadden, uitgesloten.

De huisarts hield veel patiënten zelf onder behandeling, ook als er gedacht werd aan angina pectoris. Dit is opmerkelijk, daar in de specialistische literatuur meestal wordt aanbevolen patiënten met recent ontstane angina pectoris-klachten snel te verwijzen. Een recent begin wordt namelijk beschouwd als een kenmerk van een instabiele angina pectoris.

Elektrocardiografie werd meestal gebruikt na afloop van het eerste contact, om bij niet-verwezen patiënten bij wie een cardiale oorzaak overwogen wordt de klachten nader te evalueren. Voor besluitvorming in acute situaties is het ECG van weinig betekenis. Het is niet erg betrouwbaar bij de vroege diagnostiek van het verse infarct en dient het verwijsbeleid bij klinische verdenking hierop niet te bepalen.<sup>1,5,12-14</sup>

De in tabel 2 vermelde voorspellende waarden zijn in overeenstemming met de algemene bevindingen van *Wright and Kane*, die inhouden dat in de huisartspraktijk de voorspellende waarde van een goede prognose steeds hoger is dan die van een minder goede prognose.<sup>15</sup> Dat ligt overigens voor de hand, omdat goede prognoses in de huisartspraktijk nu eenmaal veel meer voorkomen dan slechte.

Daarnaast blijft twijfel een factor van belang en men zou hieraan eveneens een voorspellende waarde kunnen toekennen. Uit diverse andere publikaties blijkt dat borstklachten, ook op de lange duur en zelfs na uitvoerige diagnostiek, lang niet altijd in hokjes zijn te vangen. Een retrospectief onderzoek in enkele huisartspraktijken leverde op dat 50 procent van de gevallen van pijn op de borst na een half jaar nog een *unproven*

etiology had.<sup>7</sup> Van een groep patiënten die wegens „hartinfarct-suggestieve“ klachten op een hartbewakingsafdeling waren opgenomen, werd 13 procent ontslagen met de diagnose *chestpain? cause*.<sup>16</sup>

Veel toestanden kunnen pijn veroorzaken die lijkt op coronaire pijn: aandoeningen van de thoraxwand, spondylose, wortelpijn, herpes zoster, hyperventilatie en met name pijn van oesophageale oorsprong.<sup>17</sup> Het diagnostiseren of het zien aankomen van een infarct blijkt dan ook niet gemakkelijk. Dit kwam al eerder naar voren in de IMIR-studie: door de huisartsen was aanvankelijk de helft van de infarcten niet herkend.<sup>5</sup> Ook specialisten kennen dit probleem: volgens een recente publikatie waren van 100 door middel van autopsie bewezen infarct-gevallen er slechts 53 tevooren als zodanig gediagnostiseerd, ondanks het gebruik van allerlei diagnostische faciliteiten.<sup>18</sup>

De ontdekking of het vermoeden van een hartinfarct leidt niet altijd tot verwijzing. In de IMIR-studie werd bijna de helft van de van een infarct verdachte patiënten niet direct verwezen.<sup>19</sup> Het is inmiddels een breed aanvaarde opvatting dat patiënten met een oncompliceerd infarct dat reeds meer dan enige uren oud is, in principe niet opgenomen hoeven te worden.<sup>20,22</sup>

We weten dat de specialist te maken krijgt met een geselecteerde groep patiënten met klachten op de borst, maar in feite geldt hetzelfde voor de huisarts: *Huygen e.a.* vonden dat slechts 10 procent van de mensen met pijn op de borst daarvoor de huisarts raadpleegde.<sup>21</sup> Het is nog maar de vraag of — en zo ja, hoe — meer diagnostiek moet worden toegepast op meer mensen. Volgens *Todd* kunnen alle moderne tests tezamen een klein infarct niet met zekerheid uitsluiten.<sup>13</sup>

Overgaan tot verregaande diagnostiek of juist een terughoudende opstelling bij specifieke borstklachten: de zin van beide beslissingen verdient evenveel studie en overweging. Omgaan met twijfels en onzekerheden is daarbij niet te omzeilen.

<sup>1</sup> Pool J, Van der Does E, Lubsen J. De diagnostiek van het dreigende hartinfarct in de huisartsenpraktijk. Ned Tijdschr Geneesk 1979; 123: 1457-60.

<sup>2</sup> Pool J, Van der Does E, Lubsen J. De diagnostiek van het dreigende hartinfarct in de huisartsenpraktijk. Ned Tijdschr Geneesk. 1979; 123: 1505-8.

<sup>3</sup> Van Ree JW. De ECG-werkgroep Nijme-

gen. Huisarts en Wetenschap. 1983; 126: 29-132.

<sup>4</sup> Tillema W, Holtkamp HA. Het cardiofonie project Zwolle. Huisarts en Wetenschap 1983; 26: 132-5.

<sup>5</sup> Meyhoom WA. De beperkingen van cardiofonie. Huisarts en Wetenschap 1983; 26: 136-9.

<sup>6</sup> Lamberts H. Waarom gaat iemand naar de huisarts? Een internationale studie met de Reason for Encounter Classification. Huisarts en Wetenschap 1984; 27: 234-44.

<sup>7</sup> Blalock SM. The symptom of chestpain in family practice. J Fam Pract 1977; 4: 429-33.

<sup>8</sup> Lamberts H. Incidentie en prevalentie van gezondheidsproblemen in de huisartsenpraktijk. Huisarts en Wetenschap 1982; 25: 401-14.

<sup>9</sup> Continue Morbiditeitsregistratie NUH 1971-1978. Werkgroep Epidemiologie in de huisartspraktijk. Gewone ziekten. Nijmegen: Nijmeegs Universitair Huisartsen instituut, 1980.

<sup>10</sup> Goldman L, Weinberg M, Weisberg M, et al. A computer-derived protocol to aid in the diagnosis of emergency room patients with acute chestpain. N Eng J Med 1984; 310: 1273-8.

<sup>11</sup> Silverman KJ, Grossman W. Angina pectoris, natural history and strategies for evaluation and management. N Eng J Med 1984; 310: 1712-7.

<sup>12</sup> Smit PTh. Spoedgevallen in een huisartsenpraktijk. 3e dr. Utrecht: Bohn, Scheltema de Holkema, 1976.

<sup>13</sup> Todd JW. Query cardiac pain. Lancet 1983; i: 330-2.

<sup>14</sup> De Dombal ET. Evaluation of decision making by humans and computer in acute abdominal and acute chest pain. Lecture notes in medical Informatics 1983; 22: 42-53.

<sup>15</sup> Wright DD, Kane RL. Predicting the outcome of primary care. Med Care 1982; 20: 180-7.

<sup>16</sup> Wilcox RG, Roland JM, Hampton JR. Prognosis of patients with „chest pain? cause“. Br Med J 1981; 282: 431-3.

<sup>17</sup> Bennet JR. Chest pain: heart of gullet? Br Med J 1983; 286: 1231-2.

<sup>18</sup> Zarling EJ, Sexton H, Milnor P. Failure to diagnose acute myocardial infarction. JAMA. 1983; 250: 1171-81.

<sup>19</sup> Van der Does E, Lubsen J. Acute coronary events in general practice [Dissertatie]. Rotterdam: Erasmus Universiteit Rotterdam, 1978.

<sup>20</sup> Eggertsen SC, Berg AO. Is it good practice to treat patients with uncomplicated myocardial infarction at home? JAMA 1984; 251: 349-50.

<sup>21</sup> Jones R. Myocardial infarction and the general practitioner. Practitioner, 1984; 228: 377-84.

<sup>22</sup> De Melker RA. Thuisbehandeling myocardinfarct? Practitioner (Nederlandse uitgave) 1984; 1: 959-61.

<sup>23</sup> Huygen FJA, Van den Hoogen H, Neefs WJ. Gezondheid en ziekte; een onderzoek van gezinnen. Ned Tijdschr Geneesk 1983; 127: 1612-9.



### DEEL III

#### MEDISCHE BESLISKUNDE, CONCEPTEN EN TOEPASSINGEN.

## Principes van besliskunde

*J. A. Knottnerus\**

### 1 Inleiding

Beslissingen moeten juist daar worden genomen waar sprake is van onzekerheden. Voor de huisarts zijn deze aan de orde van de dag. Hij heeft te maken met vele niet uitgekristalliseerde hulpvragen en problemen en de zich aan hem presenterende ziektebeelden verkeren vaak in nog weinig herkenbare stadia. In bijna ieder arts-patiëntcontact moeten dan ook keuzen worden gemaakt uit alternatieven. Het streven daarbij is door te zeilen tussen de Scylla van het missen van belangrijke diagnoses en behandelingsmogelijkheden en de Charybdis van de somatische fixatie en de iatrogenese. De huisarts, die immers de toegangspoort tot de geneeskunde markeert, heeft hierbij een bijzondere verantwoordelijkheid. Hiermee zij aangegeven dat huisartsgeneeskunde voor een belangrijk deel besliskunde is.

De bedoeling van dit hoofdstuk is om vanuit de nog jonge medische besliskunde enkele basale principes aan te dragen, die van nut kunnen zijn bij het bestuderen en verbeteren van de huisartsgeneeskundige besluitvorming.

### 2 Hypothesevorming

Het meest richtgevende gebeuren in de diagnostiek van de arts is misschien wel de vorming van hypothesen omtrent het gezondheidsprobleem of ziektebeeld dat de patiënt mogelijk parten speelt. Meestal worden deze hypothesen al vroeg in het consult – soms na enkele seconden – en op grond van luttele informatie opgesteld.<sup>1</sup> In de loop van het verdere consult probeert de arts nu door middel van anamnese, onderzoek en eventuele nadere diagnostiek de hypothesen te bevestigen of te verwerpen. Telkens vraagt hij zich af: welke bevindingen zouden zich voordoen als de hypothese waar is? En: hoe waarschijnlijk wordt de hypothese, gegeven deze bevindingen? Deze strategie transformeert de vrijwel onbegrensde, open vraagstelling: wat is er aan de hand met deze patiënt? tot één of meer gesloten vraagstellingen. Is er sprake van galstenen? Ulcus duodeni? Cholecystitis? Of is de patiënt overwerkt? Zo'n set van alternatieven kan getest worden. De arts kan van elk van de hy-

---

\* Arts, Capaciteitsgroep Epidemiologie en Gezondheidszorgonderzoek, Rijksuniversiteit Limburg

pothetische eindpunten toewerken naar datgene wat hij bij de patiënt vindt.

De hypothesen, en hun volgorde van belangrijkheid, worden in het algemeen gebaseerd op een beoordeling van:

- a waarschijnlijkheden: op welke aandoeningen of gezondheidsproblemen heeft deze patiënt met deze klacht een verhoogde kans, en in welke mate is deze kans verhoogd?
- b de mogelijke ernst: zijn de mogelijke aandoeningen dermate ernstig dat meer zekerheid erover gewenst is;
- c de therapeutische consequenties: heeft het stellen van de mogelijke diagnose zodanige consequenties voor het therapeutisch beleid, dat zij per se moet worden uitgesloten of bevestigd?

Aan de fase van de vroege hypothesevorming naar aanleiding van een bepaalde klacht wordt in de medische opleiding weinig aandacht besteed. Het betreft echter een cruciale fase in de diagnostiek en fouten die in deze fase worden gemaakt, worden niet gemakkelijk gecorrigeerd.

Zo blijkt dat een eenmaal gevormde hypothese door artsen niet snel wordt verlaten, ook al is zij onjuist en niet in overeenstemming met latere bevindingen. Soms kan men vastraken op een fout spoor: als men eenmaal aan angina pectoris denkt, is het niet gemakkelijk de diagnose hyperventilatie te stellen, en andersom. Een tweede probleem is, dat het testen van een hypothese zelden een 100%-zeker antwoord kan opleveren. Er blijft een bepaalde mate van onzekerheid over en die onzekerheid, hoe gering ook, kan een probleem op zich worden of alsnog aanleiding geven tot onnodig handelen en iatrogenese: 'dokter, weet u zeker dat ik geen kanker heb?'

Men moet dus goede argumenten hebben om een hypothese te vormen en vervolgens te gaan testen. Dan duikt echter een derde probleem op: het is, zeker op het niveau van de huisarts niet altijd bekend of, en in welke mate, bepaalde klachten samenhangen met bepaalde aandoeningen. Met andere woorden: in hoeverre een klacht tot één of meer hypothesen zou moeten leiden is niet duidelijk.

Een voorbeeld: het is een wijdverbreid gebruik om in geval van rugklachten op een gegeven moment een röntgenfoto te maken van de lumbale wervelkolom. Uit onderzoek komt echter naar voren dat er geen duidelijk verband is tussen lage rugklachten en röntgenologische bevindingen. Zodra men dus bij iemand met lage rugklachten de hypothese van een rugafwijking gaat toetsen door middel van een foto, loopt men een fikse kans (bij ouderen is die kans tientallen procenten) dat men een 'afwijking' vindt die niets met de klachten te maken heeft.<sup>2,3</sup> De ten onrechte bevestigde hypothese zal dan moeilijk te corrigeren zijn en daarmee ligt voor de onderzochte een verdere carrière als rugpatiënt open. Wetenschappelijk onderzoek naar de inhoudelijke aspecten van de hypothesevorming is al met al van groot belang. Het is dan ook verheugend, dat er sinds de eerste aanzetten van Hodgkin<sup>4</sup> en Morell<sup>5</sup> steeds meer klachtgericht onderzoek vanuit de huisartspraktijk geëntamerd wordt. Meer dan tot nu toe zouden daarbij vergelijkingen gemaakt moeten worden ten opzichte van personen die de betreffende klacht niet hebben. Immers, alleen als



mensen met een klacht meer kans op een bepaalde aandoening hebben dan vergelijkbare personen zonder die klacht, heeft de klacht diagnostische betekenis.

### 3 Interpretatie van diagnostische gegevens

Als de huisarts door middel van anamnese en onderzoek hypothesen gaat testen, probeert hij in feite de vooraf ingeschatte kansen te herzien. Bestaande onzekerheden dienen hierdoor te worden verkleind. Dit geeft al direct aan, dat diagnostische procedures (tests) die niet tot een herziening van kansen leiden niet zinvol, en onnodig belastend zijn.

Voor het beoordelen van de vraag in hoeverre een diagnostisch gegeven informatief is, alsmede voor het kritisch interpreteren van zo'n gegeven, is het van belang enkele basisbegrippen te kennen. We bespreken deze aan de hand van een zogenaamde viervelden tabel of beslissingsmatrix (figuur 1).

*Figuur 1*

*A priori kansen, sensitiviteit en specificiteit, alsmede voorspellende waarden, zoals gedemonstreerd aan de hand van de vierveldentabel.*

	<i>ziek</i>	<i>niet-ziek</i>	<i>totaal</i>
<i>afwijkende uitslag</i>	<i>TP</i>	<i>FP</i>	<i>TP + FP</i>
<i>niet-afwijkende uitslag</i>	<i>FN</i>	<i>TN</i>	<i>FN + TN</i>
<i>totaal</i>	<i>TP + FN</i>	<i>FP + TN</i>	<i>TP + FP + TN + FN</i>

$$- \text{ a priori kans op de ziekte} = \frac{TP + FN}{TP + FP + TN + FN}$$

$$- \text{ sensitiviteit van de test} = \frac{TP}{TP + FN}$$

$$- \text{ specificiteit van de test} = \frac{TN}{TN + FP}$$

- *a posteriori-kansen*

$$\bullet \text{ voorspellende waarde van afwijkende uitslag (VW+)} = \frac{TP}{TP + FP}$$

$$\bullet \text{ voorspellende waarde van niet-afwijkende uitslag (VW-)} = \frac{TN}{TN + FN}$$

3-1 *De 'a priori' kans:* de kans op een bepaalde aandoening vóórdat er nader onderzoek is gedaan. Men kan deze kans opvatten als het percentage personen dat die ziekte zou hebben in een populatie van personen met dezelfde kenmerken en klachten als de patiënt die men voor zich heeft. De a priori kans komt dus overeen met de prevalentie in een dergelijke hypothetische populatie. Soms is over deze kans vrij veel bekend, vaak kan zij slechts worden geschat.

Het schatten van a priori kansen is een hoeksteen van het vormen van diagnostische hypothesen.

3-2 *De sensitiviteit van de test:* de kans dat in aanwezigheid van de ziekte de test ook inderdaad een afwijkende uitslag oplevert. Als men de test zou doen op een groot aantal zieken, dan geldt:

$$\text{sensitiviteit} = \frac{\text{terecht positieve uitslagen (TP)}}{\text{terecht positieve (TP) + fout negatieve (FN) uitslagen}}$$

3-3 *De specificiteit van de test:* de kans dat in afwezigheid van de ziekte de test ook werkelijk een niet-afwijkende uitslag oplevert. Als men de test doet op een groot aantal niet-zieken geldt:

$$\text{specificiteit} = \frac{\text{terecht negatieve uitslagen (TN)}}{\text{terecht negatieve (TN) + fout positieve (FP) uitslagen}}$$

Sensitiviteit en specificiteit bepalen te zamen de validiteit, of beter, het onderscheidend vermogen van de test. Zij hangen met elkaar samen. Deze samenhang is een functie van het zogenaamde afkappunt: dit is de cesuur tussen wat als afwijkende en niet-afwijkende uitslagen beschouwd wordt. In het algemeen geldt dat, als men door middel van een andere keuze van het afkappunt de sensitiviteit verhoogt, de specificiteit afneemt. En, omgekeerd, wordt de sensitiviteit kleiner als men de specificiteit verhoogt.<sup>6</sup> Dit wordt duidelijk aan de hand van een voorbeeld. Door middel van het inspannings-elektrocardiogram tracht men een indruk te krijgen over de mate van ischemie van de hartspier, ten einde besluiten te nemen ten aanzien van nadere diagnostiek of therapie. Als men nu de cesuur voor wat men nog een normale ST-depressie vindt legt bij 2 in plaats van 1 mm, dan zullen meer mensen zonder ischemie ook als zodanig geclassificeerd worden (de specificiteit wordt groter). Van degenen met ischemie zullen er echter minder worden opgespoord en voor nadere diagnostiek of therapie geselecteerd (sensitiviteit wordt kleiner). Kiest men een geringe ST-depressie als afkappunt, dan wordt de sensitiviteit groter en de specificiteit kleiner. De vraag is nu: wat is het optimale afkappunt bij een test?

Deze vraag is niet eenvoudig en zeker niet eenduidig te beantwoorden. Het antwoord is in de eerste plaats afhankelijk van de subjectieve waarde die patiënt en arts toekennen aan een (fout-)positieve en een (fout-)negatieve uitslag. Als men bijvoorbeeld de consequenties van een fout-negatieve uitslag

ernstiger vindt dan die van een fout-positieve, zal men de nadruk leggen op een hoge sensitiviteit. Worden daarentegen fout-positieven ernstiger gevonden, dan is een meer specifieke test extra belangrijk. Op dit moment lijkt het erop dat fout-negatieve uitslagen in het algemeen ernstiger worden gevonden dan fout-positieve, ook als de laatste leiden tot onnodig, vaak ingrijpend vervolgonderzoek of zelfs overbodige therapie. De arts zal dan ook geneigd zijn het afkappunt niet te stringent te stellen. Deze beslissing impliceert echter een keuze voor meer fout-positieven, met als mogelijk gevolg meer iatrogenese en somatische fixatie. De optimale keuze van het afkappunt hangt ook af van de a priori kans op de aandoening waarop men onderzoekt. Wanneer deze afneemt wordt bij eenzelfde afkappunt de kans op een fout-positieve uitslag groter (zie ook 3-4). Wil men deze in toom houden, dan zal het afkappunt stringenter moeten worden gekozen. Overigens kan men testuitslagen nog op een andere wijze interpreteren dan aan de hand van enkelvoudige afkappunten. In principe is er immers sprake van een glijdende schaal: bij iedere waarde van de uitslag is aan te geven wat de kans is op het aan- respectievelijk afwezig zijn van één of diverse aandoeningen. Intuïtief werken artsen al vaak met dit principe. Zo zal men bij een zeer sterk verhoogde bloedsuiker de kans op diabetes mellitus veel hoger (soms zelfs zeker) achten dan in geval van een licht verhoogde waarde.

Het zal duidelijk zijn dat het onderscheidend vermogen van een test nihil is wanneer men bij zieken en niet-zieken dezelfde verdeling van afwijkende respectievelijk niet-afwijkende uitslagen vindt. In dat geval kan men in plaats van het doen van de test even goed een munt opgooien. Deze situatie zou wel eens vaker kunnen voorkomen, of benaderd worden, dan men denkt. Helaas is echter van veel tests (ook veel gebruikte, zoals de bloedbezinking of leverfuncties) het vermogen om te onderscheiden tussen zieken en niet-zieken op het niveau van de huisarts niet of nauwelijks onderzocht. Meestal werkt men met 'normaalwaarden' die op statistische gronden zijn vastgesteld na onderzoek van gezonde populaties.

Het onderscheidend vermogen van een test is 'perfect' als de verdelingen bij zieken en niet-zieken elkaar in het geheel niet overlappen. Deze situatie is zeldzaam, zeker bij variabelen die op een continue schaal gemeten worden (bloeddruk, allerlei laboratoriumbepalingen). Soms echter wordt zij goed benaderd, zoals bij de uitslagen van röntgenfoto's bij fracturen.

3-4 *De voorspellende waarde van de test:* hierbij gaat het om de kans dat iemand met een afwijkende uitslag ook werkelijk de betreffende aandoening heeft (voorspellende waarde van de afwijkende testuitslag,  $vw+$ ) en de kans dat iemand met een niet-afwijkende uitslag de aandoening niet heeft (voorspellende waarde van de niet-afwijkende uitslag,  $vw-$ ). Ofwel:

$$vw+ = \frac{\text{aantal zieken met een afwijkende uitslag (TP)}}{\text{alle personen met een afwijkende uitslag (TP + FP)}}$$

$$\text{en: } \text{vw} = \frac{\text{aantal niet-zieken met een niet-afwijkende uitslag (TN)}}{\text{alle personen met een niet-afwijkende uitslag (TN+FN)}}$$

De voorspellende waarden noemt men ook wel 'a posteriori kansen', dat wil zeggen de kansen op ziekte respectievelijk niet-ziekte, ná het bekend worden van de testuitslag.

Uit de vierveldentabel is gemakkelijk het verband tussen a priori kans, sensitiviteit en specificiteit van de test, alsmede de a posteriori kans af te leiden. Dit verband staat bekend als de regel van Bayes (zie figuur 2). Het principe is

*Figuur 2*

*De formule van Bayes en haar afleiding.*

	ziek	niet-ziek
testuitslag positief	$TP = \text{sensitiviteit} \times \text{a priori kans}$	$FP = (100\% - \text{specificiteit}) \times (100\% - \text{a priori kans})$
testuitslag negatief	$FN = (100\% - \text{sensitiviteit}) \times \text{a priori kans}$	$TN = \text{specificiteit} \times (100\% - \text{a priori kans})$
	$TP + FN = \text{a priori kans}$	$FP + TN = 100\% - \text{a priori kans}$

*Afleiding formule van Bayes:*

$$\begin{aligned}
 1 \quad \text{Voorspellende waarde van een positieve testuitslag} &= \frac{TP}{TP + FP} \\
 &= \frac{\text{sensitiviteit} \times \text{a priori kans}}{(\text{sensitiviteit} \times \text{a priori kans}) + (100\% - \text{specificiteit}) \times (100\% - \text{a priori kans})} \\
 2 \quad \text{De voorspellende waarde van een negatieve testuitslag} &= \frac{TN}{TN + FN} \\
 &= \frac{\text{specificiteit} \times (100\% - \text{a priori kans})}{\text{specificiteit} \times (100\% - \text{a priori kans}) + (100\% - \text{sensitiviteit}) \times \text{a priori kans}}
 \end{aligned}$$

helder: door het toevoegen van de informatie die een bepaalde testuitslag in zich draagt aan de a priori kans, verandert deze in de a posteriori kans (voorspellende waarde).

Van groot belang is het inzicht, dat de hoogte van de a priori kans in hoge mate bepalend is voor de voorspellende waarde van de uitkomst. We geven een voorbeeld.

De sensitiviteit en de specificiteit van mammografie bij de opsporing van borstkanker kan men op grond van publikaties<sup>7, 8, 9, 10</sup> stellen op respectievelijk 90% en 85%. We kunnen het onderzoek nu verrichten bij vrouwen met verschillende a priori kansen op kanker, passend bij:

- a klachtenvrije vrouwen uit de algemene bevolking;
- b vrouwen tussen de 20 en 30 jaar, die de huisarts bezoeken met een knobbeltje in de borst;<sup>11</sup>
- c vrouwen tussen de 50 en 60 jaar, die de huisarts bezoeken met een knobbeltje in de borst.<sup>11</sup>

Gemakkelijk is nu te berekenen met behulp van een vierveldentabel of de regel van Bayes hoe groot de voorspellende waarden worden. Deze zijn gegeven in tabel 1. Het blijkt dat de voorspellende waarde van een afwijkend mammogram hoog is, als ook de a priori kans op kanker hoog is. Men moet echter dan wel meer rekening houden met een lagere voorspellende waarde van een niet-afwijkende uitslag, met andere woorden: de kans op een fout-negatieve uitslag wordt groter.

Tabel 1

*Voorspellende waarden van een afwijkend (VW+) en niet-afwijkend (VW-) mammogram bij patiënten (typen) met verschillende a priori kansen op borstkanker. Rekenvoorbeeld, uitgaande van een sensitiviteit en specificiteit van resp. 90% en 85%.\**

'type patiënt'	a priori kans	VW+	VW-
1 klachtenvrij, algemene bevolking	0.2%	-1.2%	99.99%
2 20-30 jaar, met knobbeltje in de borst bij de huisarts	5%	24.0%	99.38%
3 50-60 jaar, met knobbeltje in de borst bij de huisarts	40%	80.0%	92.72%

\* Hierbij is, om didactische redenen, geen rekening gehouden met het feit dat de sensitiviteit en specificiteit in de verschillende situaties enigszins kunnen verschillen (zie blz. 11/4-13).

Als de a priori kans zeer laag is, blijft de voorspellende waarde van een afwijkende uitslag betrekkelijk laag en zijn er relatief veel fout-positieve uitslagen te verwachten. Men zou kunnen zeggen dat een testuitslag maar een beperkt deel van de informatie is die moet worden beoordeeld. Soms is een reeds bestaande zeer sterke of juist zeer geringe verdenking op grond van de mate van aanwezigheid van risicofactoren of het klinisch beeld, minstens even zwaarwegend.

Nog een voorbeeld. Bij een vrouw van 25 jaar zonder klachten wordt bij een routine-cervixuitstrijkje een Pap I gevonden. De arts zal er dan op vertrouwen dat er geen sprake is van cervixkanker. Stel nu dat de uitslag Pap I ook afgegeven wordt bij een veertigjarige vrouw met klachten van tussentijds bloedverlies en contactbloedingen. Ondanks het feit dat de testuitslag ook nu negatief is, zal de arts in dit geval minder zeker van zijn zaak zijn: hij houdt rekening met een fout-negatieve uitslag en acht kanker nog geenszins uitgesloten. De arts past nu in feite toe wat zojuist geformuleerd is: hij interpreteert

de testuitslag met in het achterhoofd de kennis dat cervixkanker bij klachten-vrije vrouwen van 25 jaar erg zeldzaam is (lage a priori kans), maar bij vrouwen van 40 jaar met de beschreven klachten veel vaker voorkomt (hogere a priori kans).

Bevolkingsonderzoeken zijn erom berucht dat zij zich meestal richten op populaties met dermate geringe a priori kansen dat er veel fout-positieve uitslagen uit de bus rollen. Op het niveau van de huisarts, op wiens spreekuur allerlei ernstige ziekten relatief zeldzaam zijn zou men, weliswaar in mindere mate, hetzelfde kunnen vrezen. Daarom is een kritische instelling ten aanzien van het formuleren van hypothesen, alsmede een goede selectie van patiënten voor nadere diagnostiek, heel belangrijk.

Daarbij geldt natuurlijk dat de vraag van de patiënt de eerste ingang is van de hulpverlening. Het kan dan soms zo uitpakken, dat de arts op aandrang van de patiënt een hypothese gaat testen die hijzelf op goede gronden nooit zou hebben gevormd. Het is dan belangrijk dat de arts zich voor het onderzoek realiseert — en dit ook aan de patiënt duidelijk maakt — dat de a priori kans erg laag is en dat een afwijkende uitkomst heel wel fout-positief kan zijn.

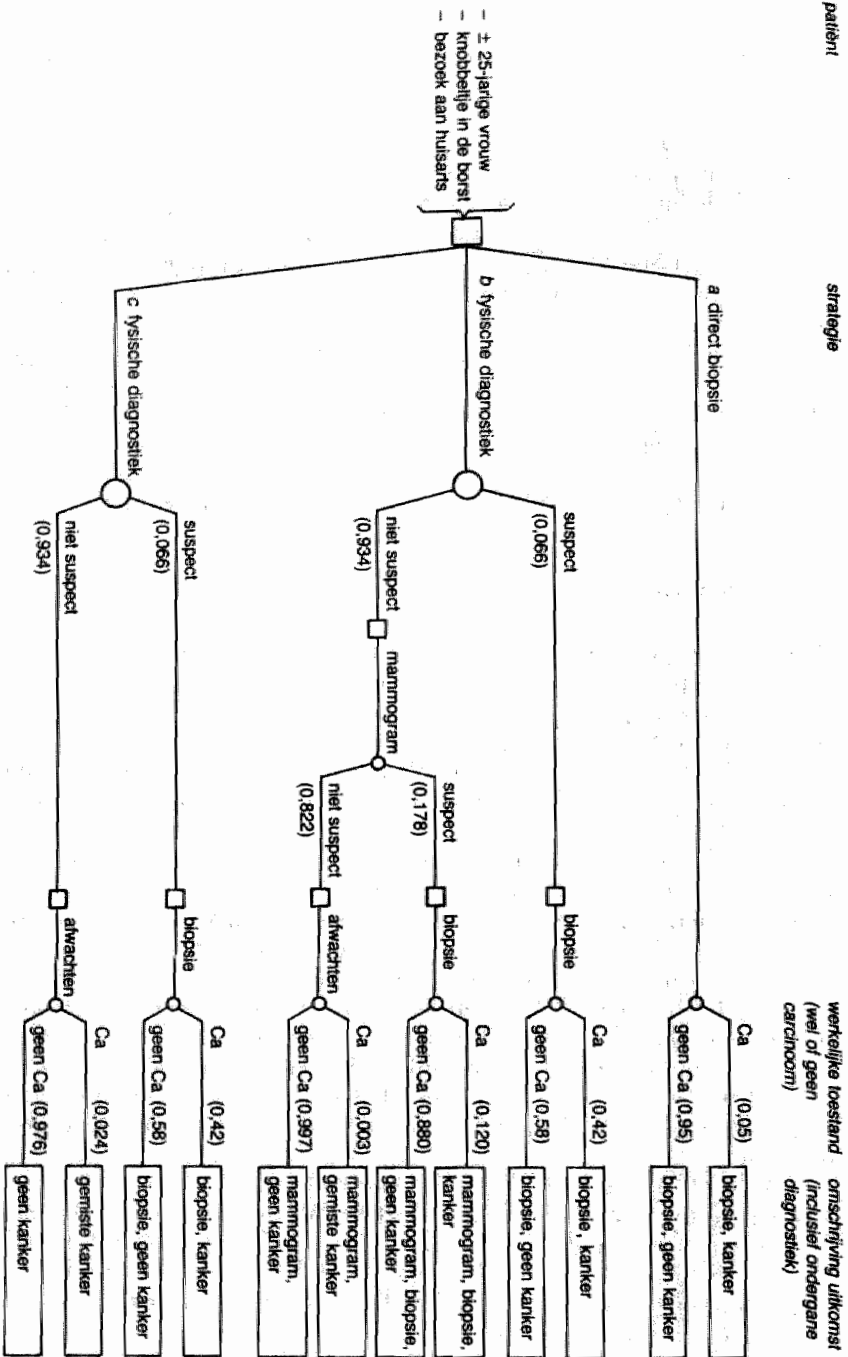
#### 4 Beslisbomen

Huisartsen moeten vaak kiezen uit verschillende diagnostische strategieën. Zij doen dit op grond van kennis, ervaring en gewoontevorming alsmede aan de hand van wat zij lezen in de vakliteratuur. Niettemin zijn er veel twijfels en verschillen van opvatting. Het is dan zinvol de diverse alternatieven zo expliciet mogelijk tegen elkaar af te wegen. Hierdoor komt men de meest wezenlijke beslispunten op het spoor en is het mogelijk witte plekken in de bestaande kennis te lokaliseren.

Eén van de manieren om dit te doen is het ontwerpen van een zogenaamde beslisboom. Daarin worden de diverse mogelijke strategieën op overzichtelijke wijze ondergebracht, te zamen met hun mogelijke uitkomsten. Een en ander wordt dan voor bestudering toegankelijk. Men kan bovendien trachten, met behulp van gegevens in de literatuur eventueel aangevuld door de mening van deskundigen, meer specifieke uitspraken te doen over de meest wenselijke optie. Daarbij moet er rekening mee worden gehouden, dat onder verschillende omstandigheden andere opties de voorkeur kunnen verdienen.

We demonstreren een en ander aan de hand van het voorbeeld over het mammaknobbelkje. Men kan staan voor de keuze uit de volgende drie alternatieven:

- a directe verwijzing naar de chirurg voor biopsie;
- b eerst wordt fysische diagnostiek verricht. Als deze een voor maligniteit suspecte uitslag oplevert wordt verwezen voor biopsie. Indien het resultaat niet suspect is wordt mammografie verricht. Is deze afwijkend, dan wordt alsnog verwezen. Anders wordt afgewacht en na enige weken een controle afgesproken;



**Figuur 3**  
Beslisboom betreffende drie verschillende diagnostische strategieën bij een jonge vrouw die de huisarts bezoekt met een knobbelje in de borst (zie voor toelichting de tekst).

*c* er wordt fysische diagnostiek verricht. Als het resultaat suspect is, wordt verwezen. In de overige gevallen wordt afgewacht en een nacontrole afgesproken.

In figuur 3 wordt een beslissing aangegeven met een vierkantje (keuze-knoop) en een uitkomst als een rondje (kansknoop). De laatste vertakking betreft de werkelijke situatie wat betreft de aanwezigheid van kanker. We gaan ervan uit dat biopsie perfecte informatie oplevert. Geheel rechts zijn de einduitkomsten per tak omschreven.

Met behulp van literatuurgegevens<sup>7, 8, 9, 10, 12</sup> over de sensitiviteit en specificiteit van fysische diagnostiek en mammografie alsmede over de kans op borstkanker bij een jonge vrouw met een knobbel in de borst<sup>11</sup> is het mogelijk de beslisboom van getallen te voorzien. In de beslisboom hebben we in plaats van sensitiviteit en specificiteit de voorspellende waarden van de diverse onderzoeksuitslagen weergegeven. Zij zijn bij een gegeven *a priori* kans uit elkaar te berekenen door middel van de formule van Bayes (figuur 2). Wij kozen voor deze weergave omdat zij meer logisch aansluit op de volgorde van het medisch besluitvormingsproces. In dit voorbeeld hebben we alleen gerekend met de gemiddelden van de literatuurgegevens. Op de technieken die ontwikkeld zijn om ook de extremen in de analyse (zogenaamde sensitiviteitsanalyse) te betrekken, kunnen we hier niet ingaan (zie hiervoor Weinstein<sup>13</sup>).

Bij iedere kansknoop is de som van de mogelijke uitkomsten 100% (1,0). Men kan de kansverdeling der verschillende uitkomsten vinden door op weg naar elke uitkomst de achtereenvolgende kansen met elkaar te vermenigvuldigen. Zo kan men nagaan dat de kans dat men bij strategie *b* een carcinoom in eerste instantie mist, gelijk is aan  $0,934 \times 0,822 \times 0,003 = 0,002$ . (= 2 van de 50 kankergevallen). Daartegenover staat dat wanneer men 1000 van deze patiënten zou onderzoeken, men bij strategie *b*  $0,934 \times 0,822 \times 1000 = 768$  directe verwijzingen voor biopsie zou uitsparen, ten opzichte van strategie *a*.

Indien men zich tot strategie *c* zou beperken dan zullen op 1000 onderzochten  $0,934 \times 0,024 \times 1000 = 22$  van de in totaal 50 kankergevallen worden gemist. Weinigen zullen strategie *c* nog acceptabel vinden.

Het is een kwestie van afweging welke keuze men uiteindelijk maakt. Onder meer hangt dit af van het gewicht dat men toekent aan het voordeel van het uitsparen van verwijzingen en biopsieën, tegenover het niet direct ontdekken van een klein aantal carcinomen.

Als een ander type patiënt onderzocht wordt kan het beeld flink veranderen. Dit komt met name omdat zoals we al eerder zagen, de voorspellende waarden van testuitslagen afhangen van de *a priori* kans op ziekte. Zo kan men nagaan dat als men strategie *b* zou kiezen voor vrouwen boven de vijftig die op het spreekuur komen met een knobbel in de borst, men veel meer carcinomen zou missen en bovendien veel minder biopsieën zou uitsparen. Strategie *a* lijkt voor deze vrouwen dus beter. Anderzijds zal men bij klachten-vrije jonge vrouwen bij toepassing van strategie *b*, als vorm van screening, waarschijnlijk relatief zeer veel biopsieën moeten doen om een sporadisch kankergeval te vinden.



Als de biopsie niet als perfecte test beschouwd wordt, maar – wat niet ir-  
reëel is – als een test die ook foute uitslagen kan opleveren, zou de keuze nog  
moeilijker worden. Men moet dan niet alleen meer rekening houden met, ach-  
teraf, onnodige verwijzingen, maar ook met onnodige operaties en anderzijds  
met ook door biopsie gemiste carcinomen.

## 5 Utiliteiten

In de vorige paragraaf is er al op gewezen dat de keuze uit verschillende stra-  
tegieën mede afhangt van de waarden die men aan de mogelijke uitkomsten  
van elk der strategieën hecht. Deze waarden noemt men ook wel 'utiliteiten'.  
Voor het doorrekenen van beslisbomen is het van belang dat deze utiliteiten  
in getalsvorm worden uitgedrukt. Soms kan men volstaan met overlevings-  
percentages. In het algemeen, zeker op het niveau van de huisarts, ligt het  
echter veel ingewikkelder. Men moet vragen beantwoorden als:

- welk risico om een appendicitis acuta te missen wil ik nog lopen, ten  
einde overbodige buikoperaties te voorkomen?
- hoe belastend is het voor een patiënt die uiteindelijk een hyperventila-  
tiesyndroom blijkt te hebben, dat hij bij de cardioloog allerlei diagnostiek  
heeft moeten ondergaan? En hoe verhoudt zich dit ten opzichte van het  
achterwege laten van de diagnostiek bij iemand die in feite wél angina pec-  
toris heeft?
- hoe groot is bij de behandeling van lichte hypertensie de winst van een  
wat geringere kans op hersenbloeding of hartinfarct, ten opzichte van de  
zekerheid van een levenslange antihypertensieve therapie?
- welke kans wil ik nog lopen op het missen van een endometriumcarci-  
noom als ik het aantal curettages en uterusextirpaties wil terugdringen?
- hoeveel beter moet het resultaat van operatie zijn ten opzichte van het  
natuurlijk beloop, om de belasting door en rondom een operatie te compen-  
seren?

Er zijn verschillende methoden ontwikkeld om dergelijke 'utiliteitsanaly-  
ses' uit te voeren.<sup>13</sup> Hoewel het hier toch gaat om het uiteindelijk nut van het  
medisch handelen, staan deze analyses nog in de kinderschoenen.

In de individuele patiëntencontacten van de huisarts zal het overleg met en  
de inbreng van de patiënt van doorslaggevend belang zijn. Een open gesprek  
over de alternatieven, hun voordelen en hun risico's is daarvoor een voor-  
waarde. Aan de basiskennis en -inzichten waarover de huisarts ten behoeve  
van zo'n communicatie zou moeten kunnen beschikken kan de medische be-  
sliskunde in de toekomst een steentje bijdragen.

De huisarts zal zich al vroeg in de diagnostische fase moeten bezinnen op de  
mogelijke uitkomsten. Van belang is dat hij vóór het entameren van een dia-  
gnostische procedure zich afvraagt:

- a is, of zijn, de aandoening(en) waarop ik wil onderzoeken wel op zin-  
volle wijze te behandelen?

*b* worden beleid en behandeling beïnvloed door de diagnostische informatie die ik nu wil gaan verzamelen of tot de verzameling waarvan ik anderen (bijvoorbeeld specialisten) aanzet?

*ad a* Als het resultaat van een behandeling niet beter is dan het natuurlijk beloop, kunnen diagnostische procedures beter achterwege blijven. Men kan denken aan het twijfelachtige nut van het opereren van een bij toeval ontdekte asymptomatische carotisstenose, of aan therapeutisch niet of nauwelijks te beïnvloeden maligniteiten.

*ad b* Voorbeelden zijn het verrichten van een controle-duodenoscopie tijdens of na de behandeling van een al aangetoond ulcus duodeni; of het verrichten van uitvoerige oncologische diagnostiek bij een ernstig zieke patiënt die een intensieve therapie niet meer zou willen of kunnen verdragen.

Deze principes lijken open deuren. Als men zich echter aanwent systematisch vooraf na te denken – en eventueel te overleggen met de patiënt of de specialist – over de verwachte waarde van de extra te vergaren informatie, zal men merken dat er heel wat diagnostiek achterwege kan blijven.

We merken nog op dat diagnostiek meer doelen kan hebben dan het bereiken van een beter therapeutisch resultaat, zoals het wegnemen van onzekerheid, of het geruststellen van de patiënt. Dit doet echter niet af aan de geldigheid van de zojuist genoemde twee principevragen.

## **6 Enkele verschillen tussen huisarts en specialist**

Na de vorige paragrafen kan een aantal belangrijke verschillen worden aangegeven voor de diagnostische posities waarin huisarts en specialist verkeren.

– In de populatie patiënten die de specialist bezoekt, is als gevolg van de voorselectie van de huisarts de *a priori* kans op ernstige ziekten en op aandoeningen die specialistische behandeling vergen groter dan in de patiëntenpopulatie die de huisarts ziet. Dat betekent dat de voorspellende waarde van een afwijkende bevinding bij een door de specialist verricht onderzoek altijd groter is dan de dezelfde bevinding door de huisarts, aannemend dat beiden in principe even kundig zijn. Ten aanzien van de afwezigheid van deze ziekten is de voorspellende waarde van een niet-afwijkende bevinding bij de huisarts juist groter.

– Voor de ziekten die de huisarts vaak ziet en de specialist minder vaak, geldt het omgekeerde. Voorbeelden hiervan zijn vele *minor ailments*, exanthemen bij kinderen, enz.

– Men kan stellen dat de voorspellende waarde van een afwijkende uitslag van specialistisch onderzoek relatief groot kan zijn, omdat de huisarts voorselecteert. Als de huisarts dit niet deed, zou veel overbodige diagnostiek en eventueel behandeling plaatsvinden. Met andere woorden: doordat de huisarts bij lage *a priori* kansen een grote voorspellende waarde van normale be-

vindingen heeft, kan de specialist bij de verwezen gevallen een hoge voorspelende waarde van afwijkende resultaten bereiken.

Een voorbeeld: bij de arts meldt zich een patiënte met een bultje in de hals. Het is licht drukgevoelig, glad en 0,5 cm in doorsnede. Het is duidelijk dat een dergelijke presentatie bij de huisarts een geheel andere betekenis heeft dan op het spreekuur van de oncoloog. Op grond van dezelfde bevindingen zal de eerste een ontstoken klier het meest waarschijnlijk achten en een afwachten-de houding aannemen. De laatste zal echter een primaire maligniteit of metastase willen uitsluiten met behulp van verder onderzoek. Men kan zich ook voorstellen dat wanneer beiden een klinische les over dit symptoom zouden geven, zij met sterk verschillende aanbevelingen zouden eindigen.

Ook een ander gevolg van verschillende prevalenties (en dus a priori kansen) in de populaties van huisarts en specialist moet in overweging worden genomen. Zoals gezegd hangt de optimale keuze van het afkappunt immers mede af van de prevalentie. Het is dus niet vanzelfsprekend dat in situaties waarin verschillende prevalenties — a priori kansen — aan de orde zijn (bijvoorbeeld die van huisarts, respectievelijk specialist) dezelfde afkappunten gehanteerd worden. Dit punt wordt tot nu toe echter in het algemeen niet betrokken bij het vaststellen en toepassen van zogenaamde normaalwaarden.

Een belangrijk verschilpunt in de morbiditeitspatronen die huisarts en specialist onder ogen krijgen, is voorts het ontwikkelingsstadium van het klinisch beeld waarmee aandoeningen zich manifesteren. In het algemeen zullen de verschijnselen in de fase waarmee de huisarts ermee geconfronteerd wordt, minder duidelijk ontwikkeld zijn dan in de fase na eventuele verwijzing naar de specialist. In de eerste situatie zullen de zieken minder goed te herkennen zijn en ook minder goed van niet-zieken kunnen worden onderscheiden dan in de tweede situatie. In de huisartsenfase is het onderscheidend vermogen van onderzoeksmethoden dan ook veelal lager dan in de specialistenfase. Zo is bijvoorbeeld het onderscheidend vermogen van de thermometer in de allereerste fase van appendicitis acuta nihil, hetzelfde geldt voor leverfunctietests bij beginnende of lichte hepatitis. Refererend aan het voorbeeld over borstkanker: de specialist onderzoekt gemiddeld grotere tumoren en deze kunnen beter gevoeld worden.

In de regel zal dus de sensitiviteit van een onderzoek groter zijn als de aandoeningen die men op het oog heeft verder ontwikkeld zijn. Soms is het zo dat de specificiteit echter tegelijk minder hoog wordt. Bijvoorbeeld als er relatief veel verdachte gevallen worden verwezen naar de specialist. Zo kan men zich voorstellen dat een chirurg te maken krijgt met een overmaat aan suspect aanvoelende goedaardige tumoren, die heel moeilijk door middel van palpatie van kwaadaardige tumoren zijn te onderscheiden. Algemeen gesteld: zieken in de huisartspraktijk lijken meer op gezonden (alleen de herkenbare gevallen worden verwezen) en gezonden in het ziekenhuis lijken meer op zieken (anders waren zij niet verwezen). Huisarts en specialist doen met een geheel ander spectrum van ontwikkelingsstadia klinische ervaringen op.

Al met al komt de vraag op, of onderzoeksmethoden die 'geijkt' zijn in een

specialistische populatie, in een huisartsensetting wel zonder meer volgens dezelfde criteria kunnen worden geïnterpreteerd. Recente inzichten steunen de opvatting, dat de validiteit van onderzoeksmethoden moet worden vastgesteld in een populatie die op relevante punten representatief is voor de personen waarop deze methoden zullen worden toegepast.<sup>14, 15</sup>

Tenslotte: het is mogelijk dat door bepaalde aangeleerde of ingesleten diagnostische gewoonten, verbanden tussen klachten en aandoeningen worden gesuggereerd die niet of nauwelijks bestaan. Als huisartsen en specialisten tijdens de opleiding of nascholing een bepaald diagnostisch of verwijs-gedrag krijgen ingegoten, kan dit leiden tot een 'self-fulfilling prophecy'. Daardoor kan de specialist telkens in zijn opvatting worden bevestigd dat een bepaald symptoom samenhangt met een bepaalde aandoening, terwijl dit op het niveau van de huisarts in feite niet zo is. Elders is dit mechanisme nader gedemonstreerd.<sup>16</sup> Zo bestaat er gerede twijfel of de klacht moeheid en het hemoglobinegehalte<sup>17</sup>, en ook hoofdpijn en hoge bloeddruk<sup>18</sup>, wel een relevante samenhang met elkaar vertonen. Echter, als huisartsen bij deze klachten relatief vaak het Hb respectievelijk de bloeddruk meten en als mensen met afwijkende bevindingen worden verwezen, zal de specialist een verband tussen de betreffende klachten en afwijkingen vinden.

## 7 Slot

In dit hoofdstuk zijn slechts basale notities aangereikt. Kernpunt van de medisch-besliskundige benadering is het kritisch beschouwen van de waarde van diagnostische gegevens, te zamen met het volgens een logische systematiek analyseren van het besluitvormingsproces. Hierdoor kunnen:

- de punten waarop besluiten moeten worden genomen helder worden gelokaliseerd;
- de eventuele consequenties van die besluiten voor het verdere beleid worden bestudeerd;
- diverse strategische opties tegen elkaar worden afgewogen;
- onzekerheden en manco's in de wetenschappelijke kennis worden opgespoord en voor onderzoek toegankelijk gemaakt;
- vruchtbare aanknopingspunten worden gevonden voor medisch onderwijs, nascholing en intercollegiale toetsing.

Essentieel is dat men zich steeds de vraag stelt of het besluit dat men overweegt relevante consequenties voor het beleid heeft en of de te verwachten voordelen voor de patiënt opwegen tegen de nadelen.

Het zal lang niet altijd mogelijk zijn op grond van bestaande gegevens een betrouwbare, gekwantificeerde, indruk te krijgen van de waarde van de verschillende opties. Bovendien is de werkelijkheid van elk individueel consult weer zo complex dat nooit alle relevante overwegingen in een algemene analyse kunnen worden betrokken. Een kwalitatieve structurering van het besluit-

vormingsproces, in bijvoorbeeld een beslisboom, kan echter op zichzelf al zeer verhelderend werken.

Pasklare oplossingen moet men van een beslistkundige benadering niet verwachten. Integendeel: routinematig handelen komt juist ter discussie. De patiënt, die het eerste en het laatste woord dient te spreken, kan bovendien een beter inzicht worden verschaft in de implicaties van een bepaalde aanpak.

## Literatuur

- 1 Elstein, A. S., L. S. Shulman en S. A. Sprafka, *Medical Problem Solving, an analysis of clinical reasoning*. Harvard University Press, Cambridge, Massachusetts 1978.
- 2 Valkenburg, H. A., Reumatische klachten en aandoeningen in een open Nederlandse bevolkingsgroep (4). *Reuma Wereldwijd* (1980) 4:1-7.
- 3 Sanders, H. W. A., Klinische betekenis van degeneratieve afwijkingen van de lumbale wervelkolom en consequenties van het aantonen ervan. *Nederlands Tijdschrift voor Geneeskunde* (1983) 127:1374-7.
- 4 Hodgkin, K., *Towards earlier diagnosis in primary care*. Londen 1978.
- 5 Morell, D. C., *Huisartsgeneeskunst*. Bunge, Utrecht 1978.
- 6 Sturmans, F., *Epidemiologie, theorie, methoden en toepassing*. Dekker & v.d. Vegt, Nijmegen 1982.
- 7 Dronkers, D. J., De betrouwbaarheid van de mammografie. *Nederlands Tijdschrift voor Geneeskunde* (1977) 123:955-7.
- 8 Clarijs, W. W. J., M. van Deelen-Willems, J. H. C. L. Hendriks, E. J. C. Lubbers en W. J. H. Schmidt, Het occulte mammacarcinoom. *Nederlands Tijdschrift voor Geneeskunde* (1977) 123:526-9.
- 9 Molenaar, J. C. en B. den Herder-Kroon, De praktische waarde van de mammografie voor de chirurg. *Nederlands Tijdschrift voor Geneeskunde* (1970) 114:1688-90.
- 10 Griner, P. F., R. J. Mayewski, A. I. Musklin en P. Greenland, Selection and interpretation of diagnostic tests and procedures. *Annals of Internal Medicine* (1981) 94:533-96.
- 11 Velden, H. G. M. van der, Diagnose of prognose. De betekenis van de epidemiologie voor het handelen van de huisarts. *Huisarts en Wetenschap* (1983) 26:125-8.
- 12 Rombach, J. J., *Breast cancer screening. Results and implications for diagnostic decision making*. Stafleu, Alphen aan den Rijn 1980.
- 13 Weinstein, M. C. en H. V. Fineberg, *Clinical decision analysis*. Saunders, Philadelphia 1980.
- 14 Ranschoff, D. F. en A. R. Feinstein, Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. *New England Journal of Medicine* (1978) 299:926-30.
- 15 Rümke, C. L., Kanttekeningen over de gevoeligheid, de specificiteit en de voorspelende waarden van diagnostische tests. *Nederlands Tijdschrift voor Geneeskunde* (1983) 127:556-61.
- 16 Knottnerus, J. A., Interpretatie van diagnostische gegevens. *Huisarts en Wetenschap* (1983) 26:363-8.
- 17 Knottnerus, A. en P. Knipschild, Hemoglobine en 'anemische' klachten zoals moeheid. *Tijdschrift voor Sociale Gezondheidszorg* (1983) 61:788-92.
- 18 Knipschild, P. en A. Visser, *Hoofdpijn en hoge bloeddruk*. 1984 (wordt gepubliceerd).

## Hoofdstuk 12

# De zeeffunctie van de huisarts

## Het totale plaatje is van belang

J. A. KNOTTNERUS

capaciteitsgroep gezondheidszorgonderzoek/epidemiologie, Rijksuniversiteit Limburg

### Inleiding

Afgaande op het aantal verwijskaarten dat op dit moment aan ziekenfondspatiënten wordt verstrekt, kan men de indruk krijgen, dat de 'zeef' van de huisarts niet goed functioneert: gemiddeld worden ongeveer 500 verwijskaarten per 1000 ziekenfonds-verzekerden per jaar uitgegeven.<sup>1</sup>

Over de achtergronden van dit cijfer is veel geschreven. Gewezen is op het grote beroep dat op de huisarts gedaan wordt, op de hoogge-

spannen verwachtingen die patiënten hebben van de specialistische geneeskunde, en op het groeiende aandeel van de verwijskaarten dat bijna autonoom 'wegvloeit' via langdurige poliklinische begeleiding van patiënten. Ook wordt nog al eens als vrees geuit dat een afkalving van de kennis en kunde van de huisarts gaande is, waardoor hij de problemen van patiënten niet meer zelf kan afhandelen.

Hoe dit ook zij, gelukkig is het nog steeds zo dat verreweg de meeste huisarts-patiëntcontacten niet tot verwijzing leiden. Uitgaande van 10.000 contacten per jaar in een standaardpraktijk van 2500 zielen, worden er gemiddeld zo'n 125 verwijskaarten verstrekt op 1000 contacten. Dat wil zeggen, bij minstens 87,5% van de contacten wordt niet verwezen. Minstens, omdat de verstrekking van een groot aantal verwijskaarten zich buiten de invloedssfeer van de huisarts voltrekt.

De zeeffunctie van de huis-

arts is er dus nog wel degelijk. De wijze waarop men het functioneren van de zeef beoordeelt, hangt onder meer af van de positie van waaruit men kijkt. Als men zich dit niet, of onvoldoende, realiseert, bestaat het gevaar van een verkeerde beeldvorming. De zeeffunctie kan hierdoor, ten onrechte, nog meer in discrediet geraken dan wellicht al is geschied. Wij illustreren dit met enkele soorten mogelijke beoordelingsfouten op dit gebied.

### 'De huisarts is altijd te laat'

Een groot deel van de ernstige aandoeningen die tegenwoordig veel voorkomen, ontstaat langzaam en is in het beginstadium moeilijk te onderscheiden van onschuldige aandoeningen of ongemakken, of van niet-organische problematiek. Maligne aandoeningen bij voorbeeld manifesteren zich meestal voor het eerst met een specifiek klachtenpatroon, waarachter ook allerlei levensmoeilijkheden schuil kunnen gaan.

**Samenvatting** *Het begrip artsenuitstel heeft bijna al een verwijtende betekenis gekregen. Men vergeet echter veelal dat de huisarts, uit zijn groot patiëntenaanbod, een zeer hoog percentage patiënten weet te lichten dat niet voor verdere diagnostiek en/of behandeling door de specialist in aanmerking komt. De tijd is een trouwe bondgenoot van de huisarts die deze factor vaak hanteert om zijn diagnostiek te verdiepen en niet primair om verwijzing uit te stellen.*

Daarom blijft de kans op een maligne aandoening bij zo'n klachtenpatroon nog vrij klein. De meeste huisartsen zullen derhalve niet direct het gehele diagnostische arsenaal willen opentrekken, vanwege de anders te verwachten overmatige diagnostiek en iatrogene bedreigingen. Bovendien geldt dat lang niet alle ernstige aandoeningen zich lenen voor (zinnvolle) opsporing in een zeer vroeg stadium.

De tijd wordt door huisartsen dan ook vaak als diagnosticum gebruikt. Er zijn nu een aantal mogelijkheden:

1. de meeste klachten gaan over zonder dat er veel aan gedaan hoeft te worden, dikwijls zelfs zonder dat er een diagnostische conclusie is getrokken;
2. in een aantal gevallen komt er na verloop van tijd een probleem of diagnose aan het licht, die aanleiding kan zijn tot het tijdig realiseren van een behandeling of verwijzing;
3. de klacht verdwijnt soms niet en leidt ook niet tot een duidelijke conclusie: men zou hier kunnen spreken van chronische, vage klachten;
4. in een relatief gering aantal gevallen komt na enige tijd een diagnose naar voren, waarbij men achteraf gezien graag eerder therapeutisch had willen ingrijpen.

De eerste twee situaties zullen wat de inbreng van de

huisarts betreft zonder veel complicaties afgewikkeld kunnen worden, en zullen weinig aanleiding vormen voor nadrukkelijke evaluatie achteraf. Het is immers gegaan zoals het hoort. De derde situatie leidt in het algemeen tot langdurig dokteren bij de huisarts en/of elders. Iedereen zal het er, wat betreft deze groep patiënten, over eens zijn, dat de hulpverlening niet eenvoudig is. In het vierde geval ligt het weer anders: de huisarts is nu, bij het uiteindelijk stellen van de diagnose, bijna altijd 'te laat' in de ogen van veel betrokkenen. Immers, het is achteraf vaak duidelijk dat de aandoening al veel langer bestaan moet hebben én dat het klachtenpatroon er ook bij paste. Voorbeelden zijn pancreascarcinoom en osteosaroom. Deze tumoren kunnen lange tijd schuilgaan achter de hypothetische 'psychogene buikklachten', respectievelijk '(hardnekkige) sportblessuren'. Bij veel mensen zullen deze hypothesen ook correct zijn. Als ze niet correct zijn, omdat er onverwacht sprake blijkt te zijn van één van de genoemde kankervormen, zijn deze laatste toch 'typische' gevallen: zij beginnen namelijk bijna altijd sluipend en gaan dan schuil achter andere vermeende, onschuldige, diagnoses. De indruk kan nu ontstaan dat de huisarts deze typische beelden onnodig 'gemist' heeft.

Daarenboven is, bij voor-

beeld bij maligne tumoren, vaak sprake van een min of meer exponentiële volumetoename in de tijd.<sup>2</sup> Daardoor kan de aandoening jarenlang onder de oppervlakte blijven. Is de diagnose eenmaal gesteld, dan blijkt het proces verder betrekkelijk snel te verlopen. Als men op basis van dit snelle beloop terugkijkt, kan men de huisarts licht 'onkunde' en artsenuitstel toedichten. In klinische lessen komt dan ook veelvuldig het gevaar van artsenuitstel van de zijde van de huisarts aan de orde. Men ziet dan echter alleen de helft van het plaatje.

### Het héle plaatje

Eén van de belangrijkste vaardigheden van de huisarts is het maken van onderscheid tussen patiënten die nader onderzocht, behandeld of verwezen moeten worden, én patiënten bij wie dat niet hoeft te gebeuren.

In tabel 1 is in een fictief voorbeeld weergegeven hoe dit selectieproces kan uitvallen. We zien hier dat de huisartsen 90% van de categorie ernstig zieken met klachtenpatroon X, die specialistisch behandeld moeten worden, inderdaad zonder 'tijdverlies' hebben verwezen: de 'sensitiviteit' van hun verwijzingsgedrag is 90%. Daarnaast is ook 10% van degenen met klachtenpatroon X, die géén specialistische behandeling behoeven, desondanks verwezen: de 'specificiteit' van het verwijz-

TABEL 1.

Voorbeeld: verwijsgedrag van huisartsen in relatie tot noodzaak van specialistische behandeling, bij 1000 patiënten met klachten X.

Specialistische behandeling			
	noodzakelijk	niet noodzakelijk	totaal
Direct verwezen	90	90	180
Niet of niet direct verwezen	10	810	820
	100	900	1000
- Prevalentie van de specialistisch te behandelen ziekte	$= \frac{100}{1000} = 10\%$		
- Sensitiviteit van verwijzing	$= \frac{90}{100} = 90\%$		
- Specificiteit van verwijzing	$= \frac{810}{900} = 90\%$		
- Bij 90/180 = 50% van de direct verwezen personen bleek specialistische behandeling ook noodzakelijk te zijn.			

gedrag is dan  $100\% - 10\% = 90\%$ . Gezien het feit dat de betreffende ziekte in deze groep patiënten met klachtenpatroon X in een kleine minderheid van de gevallen voorkomt (de prevalentie is laag), worden toch veel 'niet-ziekten' verwezen. Niettemin zouden een sensitiviteit en specificiteit beide van 90% een redelijke 'validiteit' van de huisarts betekenen, zeker als men deze cijfers vergelijkt met die van menige diagnostische test afzonderlijk.

Het ligt nu voor de hand dat veel aandacht van specialisten, die huisartsen feedback en bijscholing willen geven, zal uitgaan naar het vakje links onder: de tien aanvankelijk gemiste gevallen. Bij de huisartsen zal er op aange-

drongen worden zodanig te werk te gaan, dat dit vakje met zo weinig mogelijk gevallen gevuld wordt. Veel minder aandacht krijgen de 'terecht' en tijdig verwezen patiënten (zo hoort het immers), en het aantal verwezen patiënten zonder de ziekte (het wordt in het algemeen minder erg geacht dat een 'nietzieke' verwezen wordt, dan dat een zieke niet verwezen wordt).

Helemaal geen aandacht krijgt het vakje rechtsonder: degenen die terecht niet verwezen werden. Zij vallen namelijk buiten het blikveld van de specialist. Juist deze categorie echter geeft een beeld van de vervulling van een kerntaak van de huisarts: het vermijden van overbodige verwijzingen.

Het kan nu gebeuren dat de

beschouwing van het 'vakje links onder' zich gaat uiten in zeer dringende adviezen aan huisartsen: eerder doorsturen bij klachtenpatroon X en/of: reeds verwijzing overwegen bij het minder verdachte klachtenpatroon Y. In dat geval kan gebeuren wat in de tabellen 2 + 3 (zie pag. 1254) wordt getoond.

De sensitiviteit wordt inderdaad hoger (95%) en er worden dus minder zieken gemist. Omdat de verwijzdrempel is verlaagd, is echter de specificiteit kleiner geworden: meer mensen die de ziekte niet hebben, worden verwezen voor specialistisch onderzoek, en de 'zeefunctie' ten aanzien van niet-ziekten is achteruit gegaan (tabel 2). Dit kan versterkt worden door het afnemen van de prevalentie in de groep door de huisarts te beoordelen patiënten, indien nu reeds bij het minder verdachte klachtenpatroon Y verwijzing wordt overwogen (tabel 3).

Het proces kan zich voortzetten tot vrijwel geen gevallen meer worden gemist, c.q. niet meer 'te laat' worden verwezen. Een des te groter aantal niet-ziekten zal dan echter óók verwezen worden. Dit laatste zou alleen kunnen worden tegengegaan als, bij een toeneming van de 'sensitiviteit', de 'specificiteit' van de huisarts minstens gelijk zou blijven, of liever nog zou verbeteren.



TABEL 2.

Verwijsgedrag van huisartsen in relatie tot noodzaak van specialistische behandeling, bij verlagings van de verwijzingsdrempel (voorbeeld).

<i>Specialistische behandeling</i>			
	noodzakelijk	niet noodzakelijk	totaal
Direct verwezen	95	135	230
Niet (direct) verwezen	5	765	770
	100	900	1000
- Prevalentie = $\frac{100}{1000} = 10\%$ - Sensitiviteit = $\frac{95}{100} = 95\%$ - Specificiteit = $\frac{765}{900}$ - Bij 95/230 = 41% van de direct verwezen personen bleek specialistische behandeling ook noodzakelijk te zijn.			

TABEL 3.

Verwijsgedrag van huisartsen in relatie tot de noodzaak van specialistische behandeling, bij het minder verdachte klachtenpatroon Y (met lagere prevalentie) én bovendien een lagere verwijzingsdrempel (voorbeeld).

<i>Specialistische behandeling</i>			
	noodzakelijk	niet noodzakelijk	totaal
Direct verwezen	57	141	198
Niet (direct) verwezen	3	799	802
	60	940	1000
- Sensitiviteit = $\frac{57}{60} = 95\%$ - Specificiteit = $\frac{141}{940} = 85\%$ - Prevalentie = $\frac{60}{1000} = 6\%$ - Bij 57/198 = 29% van de direct verwezen personen bleek specialistische behandeling ook noodzakelijk te zijn.			

## Conclusie

Voor een juiste waardering van de zeeffunctie van de huisarts is het belangrijk dat:

Alle vier velden (zie de ta-

hellen) van de huisartsge-neeskundige betrokkenheid bij de besluitvorming ten aanzien van verwijzing overzien worden.

Inzicht bestaat in de voor-spellende waarde van klachten(-patronen) in de huisartspraktijk, en dus in de prevalentie van aandoeningen uitgaande van een bepaalde symptomatologie.

- De diagnostiek in de huis-artspraktijk geëvalueerd c.q. verbeterd wordt wat betreft haar onderscheidend vermogen. Hierbij is dan ook het pakket aan diagnostische faciliteiten van de huisarts en diens mogelijkheden tot optimale interpretatie van onderzoek-resultaten aan de orde.

Deze punten onderstrepen het grote belang van wetenschappelijk onderzoek in de huisartspraktijk voor het op peil houden, of zo men wil op peil brengen, van de zeef-functie van de huisarts. □

## Literatuur

- (1) C.B.S., *Handboek voor gezondheidsstatistiek*, Nederland, 1981, Staatsuitgeverij 's-Gravenhage, 1981.
- (2) Reinhold, H. S. (red.) *Tumorbioologie en radiologie voor klinici*, Integraal Kankercentrum Rotterdam, 1984.
- (3) De Melker, R. A.: *De selecterende zeef-functie van de huisarts: fact of fiction?* Editorial The Practitioner, Ned. editie, 1985; 2, 6: 544-545.

# Diagnostiek van de palpabele lymfeklier

door huisarts en specialist, een besliskundige benadering

J.A. KNOTTNERUS<sup>1</sup>, J. BROEKMAN<sup>2</sup> EN Prof. Dr. F. STURMANS<sup>1</sup>

<sup>1</sup> medewerkers capaciteitsgroep Gezondheidszorgonderzoek/Epidemiologie, Rijksuniversiteit Limburg

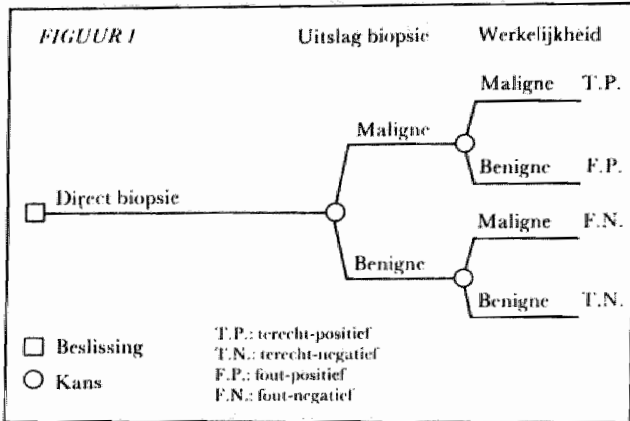
<sup>2</sup> medewerker capaciteitsgroep Pathologie, Rijksuniversiteit Limburg

## Inleiding

Over het diagnostisch beleid dat men zou moeten volgen indien door de patiënt of de arts een 'klier' wordt ontdekt, wordt verschillend gedacht. Sommigen menen dat een afwachten-de houding verantwoord is. Zij zijn van oordeel dat er zelden sprake is van ernstige pathologie.<sup>1</sup> Anderen propageren een meer agressieve aanpak, waarbij in korte tijd de gehele scala van mogelijke onderzoeken doorlopen moet worden tot de oorzaak gevonden is.<sup>2</sup> De vrees een calamiteit, met name een maligne lymfoom of een kliermetastase, over het hoofd te zien, speelt hierbij een grote rol. Een probleem is dat, voor zover onderzocht, fysisch-diagnostische be-

vindingen in het algemeen een gering onderscheidend vermogen hebben,<sup>2</sup> zeker als het gaat om de vroege ziektestadia waarmee de huisarts te maken krijgt. Ook het klachtenpatroon van de patiënt biedt vaak geen aanknopingspunt.<sup>3</sup> De belangrijkste indicatoren lijken nog de lokalisatie van de klier, en de leeftijd van de patiënt. Tot nu toe, berust helaas, vrijwel al het onderzoek dat gepubliceerd is, op gegevens die verzameld zijn op basis van diagnostische lymfeklierbiopsies bij via verwijzing geselecteerde populaties. Er wordt in deze publicaties bij lymfadenopathie dan ook gemiddeld in ongeveer 50% van de gevallen een maligne aandoening gevonden (en wel een derde hiervan primaire lymfomen, twee derden metastasen). Op grond hiervan kan de suggestie ontstaan dat in het algemeen biopsie aangewezen is. Echter, genoemde publicaties zeggen niets over de situatie van de huisarts. De kans dat een lymfeklierzwelling kwaadaardig is, is in de huisartspraktijk, afgaand op dit onderzoek, waarschijnlijk tientallen malen lager dan in de specialistische praktijk. De overgrote meerderheid van de klierzwellingen (70%) is van infectieuze oorsprong, waarbij bijzondere infecties (mononucleosis in-

**Samenvatting** *Geplaatst tegenover het probleem van de vergrote lymfeklieren zal de huisarts zich primair afvragen welke onderzoeksmethoden hij moet toepassen om het benigne of maligne karakter van de zwelling vast te stellen. Op grond van de resultaten van zijn overwegingen zal hij moeten besluiten of hij de patiënt al of niet voor verder onderzoek naar de specialist moet verwijzen. Er zijn een viertal strategieën mogelijk die de huisarts bij deze beslissingskeuze kunnen helpen.*



*Figuur 1. Iedere lymfeklierzwelling dient direct gebiopsieerd te worden.*

fectiosa, tuberculose) slechts zelden voorkomen.<sup>1</sup> Zowel bij de huisarts als bij de specialist wordt in 30 à 40% van de gevallen geen definitieve diagnose gesteld. Op grond van een eigen proefregistratie gedurende een maand in negen huisartspraktijken is te verwachten dat een huisarts met een gemiddelde praktijkgrootte per jaar 15 tot 20 maal met het probleem 'onverklaarde klierzwelling' wordt geconfronteerd. Volgens de gegevens van Lamberts<sup>5</sup> doet dit probleem zich in een normpraktijk van 2500 zielen 10 à 15 maal per jaar voor.

### Diagnostische strategieën

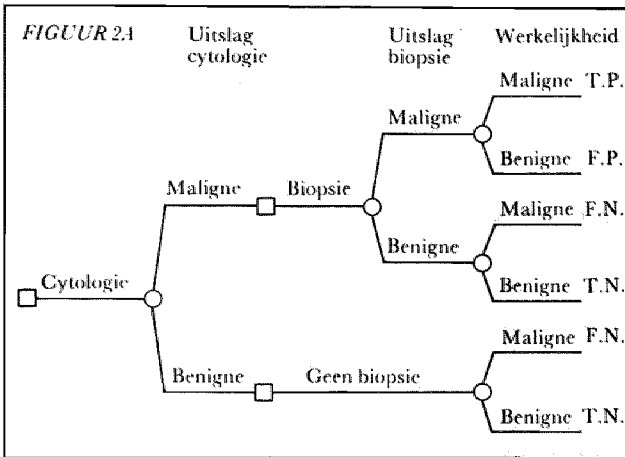
Wij onderzochten de consequenties van diverse diagnostische strategieën, zoals die te verwachten zijn op grond van gepubliceerde gegevens voor de huisarts, respectievelijk de specialist. Daarbij hebben wij ons in hoofdzaak

beperkt tot het onderscheid tussen maligne en niet-maligne, omdat dit de voor de huisarts meest wezenlijke vraagstelling betreft: 'Moet ik wel of niet verwijzen'? Binnen de groep 'maligne' dienen door de specialist dan vooral weer onderscheiden te worden primair lymfoom en metastasen van carcinoom. Wij hebben als eindpunt genomen de al dan niet juiste diagnose 'maligne of benigne'. De therapeutische consequenties hebben we in dit artikel buiten beschouwing gelaten.

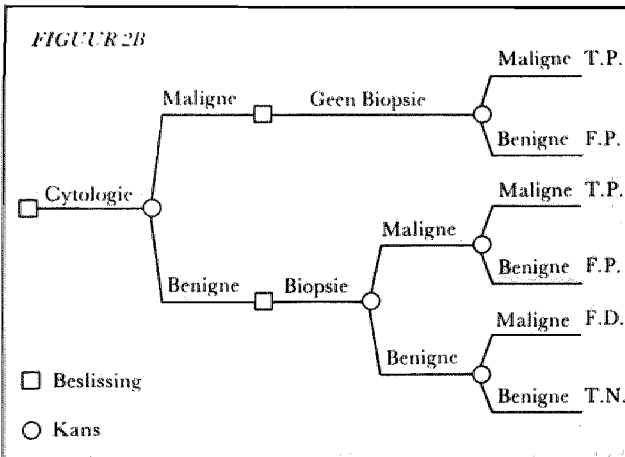
#### - Strategie 1: *Iedere lymfeklierzwelling dient gebiopsieerd te worden*

We gaan er hierbij van uit dat biopsie steeds de juiste diagnose oplevert, d.w.z. dat sensitiviteit en specificiteit steeds 100% zijn. Dit is overigens een voor biopsie gunstige veronderstelling, daar de resultaten van een aantal vervolgonderzoeken opleverden dat zich bij circa 25% van de negatieve biopsieën later alsnog een maligne lymfoom openbaarde.<sup>6, 7, 8, 9</sup> Waarschijnlijk is dit een gevolg van 'sampling error': de biopsie troffen niet de pathologische lokalisaties. Voorts zijn we uitgegaan van het optreden van (niet-fatale) complicaties in 2% van de biopsieën. De tak van de beslissingsboom, wat betreft strategie 1, is weergegeven in figuur 1.

#### - Strategie 2: *Bij iedere lymfeklierzwelling wordt eerst cytologische punctie verricht*



Figuur 2a. Eerst wordt cytologische punctie verricht. Bij positieve uitslag volgt biopsie.



Figuur 2b. Eerst wordt cytologische punctie verricht. Alleen bij negatieve uitslag volgt nog biopsie.

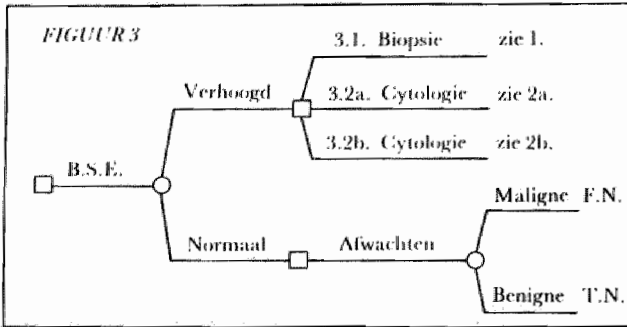
Vervolgens kan gekozen worden tussen de volgende strategieën (fig. 2).

2a als de cytologische punctie duidt op een maligne aandoening, wordt vervolgens een biopsie gedaan. Voor de huisarts betekent dit een verwijzing en vaak een opname (excisiebiopsie). Bij een negatieve cytologie wordt afgewacht;

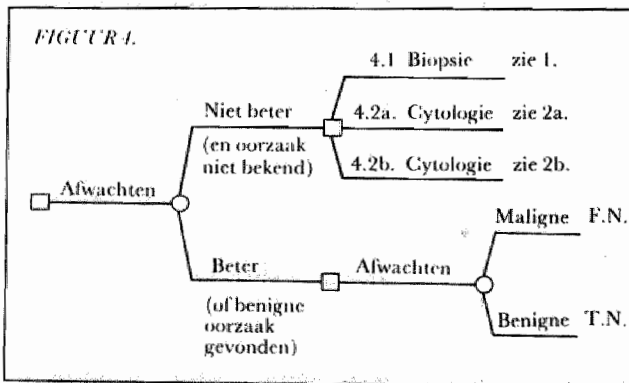
2b als de cytologische punctie negatief is, c.q. niet op maligne tekenen wijst, wordt voor de zekerheid een biopsie gedaan teneinde geen maligne aandoening te missen. Bij positieve cytologie wordt in principe gehandeld op basis van de uitslag hiervan, en wordt biopsie 'overgeslagen'. Het 'overslaan' van een biopsie kan van belang zijn, omdat een aantal ingrijpende K.N.O.-operaties in verband met aanwezigheid van maligne afwijkingen (radicale halsdissectie) moeilijker uitvoerbaar wordt na voorafgaande klierbiopsie.

Het zal niet altijd lukken om biopsie over te slaan, nl. indien cytologie wel wijst op maligne degeneratie doch geen uitsluitend geeft over het type. De strategie wordt dan minder efficiënt, c.q. er vinden meer biopsieën plaats. In de analyse gaan we uit van het gunstigste geval voor deze strategie: een positieve cytologie maakt een biopsie overbodig.

Het is duidelijk dat strategie 2b



Figuur 3. Eerst wordt een bezinking (B.S.E.) verricht. Alleen bij een te hoge uitslag wordt nadere diagnostiek verricht.



Figuur 4. Eerst wordt enkele weken afgewacht. Indien de toestand niet verbeterd is en de oorzaak nog onbekend, wordt nadere diagnostiek verricht.

voor de huisarts steeds verwijzing zal impliceren.

Er zijn, vooral voor de specialist, meer variaties in strategieën naar aanleiding van cytologie mogelijk. We beperken ons echter tot 2a en 2b, omdat deze voor het verwijzbeleid

van de huisarts van belang zijn. We zijn er vanuit gegaan, dat de sensitiviteit van cytologie voor een maligne aandoening 96% is en de specificiteit 98%.<sup>1</sup> Complicaties treden op bij 0,9% van de puncties.<sup>10</sup>

#### - Strategie 3:

*Bij iedere patiënt met een klierzwelling wordt een bloedbezinking (B.S.E.) geprikt*

Alleen indien deze te hoog uitvalt, wordt nadere diagnostiek verricht. Daarbij kan dan weer gekozen worden tussen strategie 1, 2a en 2b (fig. 3).

Op grond van literatuur nemen we aan dat de sensitiviteit van de B.S.E. ten aanzien van maligne afwijkingen gemiddeld 64%, en de specificiteit 55% is, als we ervan uitgaan dat een waarde boven de 20 mm na één uur te hoog wordt gevonden.<sup>11</sup>

#### - Strategie 4:

Een afwachterende houding wordt aangenomen. Desgewenst wordt een keelkweek, (streptokokken) en/of serologie (in het bijzonder mononucleose, cytomegalie, toxoplasmose) verricht. Indien na circa drie weken geen infectieuze oorzaak ontdekt is en de klier nog niet verdwenen c.q. kleiner geworden is, wordt de kans op een maligne aandoening uiteraard groter. Er wordt dan gekozen tussen strategie 1, 2a en 2b (fig. 4). Op grond van literatuurgegevens maakten wij een schatting van de kans dat al-

TABEL 1.

Gehanteerde kansen en veronderstellingen

<i>Prevalentie</i> van maligne afwijkingen, gegeven een lymfeklierzwelling: = a priori-kans	huisarts: 1% specialist: 50%
<i>sensitiviteit</i> t.a.v. voorkomen van maligne afwijkingen	biopsie: 100% cytologie: 96% B.S.E.: 61%
<i>specificiteit</i> t.a.v. de afwezigheid van maligne afwijkingen	biopsie: 100% cytologie: 98% B.S.E.: 55%
<i>complicaties</i>	biopsie: 2% cytologie: 0,9%
<i>Na drie weken afwachten</i>	
a specialist: - 30% van de lymfeklierzwellingen is genezen, of (terecht) als benigne gediagnostiseerd (lab.) - alle maligne klierzwellingen nog aanwezig → 'a priori-kans' = 71% ( $10/7 \times 50\%$ )	
b huisarts: - 90% van de lymfeklierzwellingen is genezen, of (terecht) als benigne gediagnostiseerd (lab.) - alle maligne klierzwellingen nog aanwezig → 'a priori-kans' = 10% ( $10/1 \times 1\%$ )	

dus na drie weken nadere diagnostiek gewenst zou zijn, en van de kans op het bestaan van een maligne aandoening in de nader onderzochte groep. Wij deden dit voor zowel verwezen<sup>3, 5, 6</sup> als niet-verwezen patiënten<sup>1</sup> (tabel 1).

### Analyse

Wij berekenden de resultaten van de diverse strategieën voor de situatie van een specialist (met een gemiddelde prevalentie van maligne afwijkingen van 50% bij patiënten van middelbare leeftijd met een lymfeklierzwelling bij een patiënt van middel-

bare leeftijd) en voor een huisarts (met een overeenkomstige prevalentie van 1%<sup>4</sup>). In tabel 1 zijn de cijfers gegeven die wij bij de analyse hanteerden. Als uitkomsten werden gekozen de aantallen verrichtingen, het aantal in eerste instantie 'gemiste' maligne aandoeningen, het geschatte aantal complicaties van biopsie en cytologie en het artsuitstel (doctor's delay). Daarbij wordt steeds uitgegaan van een 'cohort' van 1000 personen met een lymfeklierzwelling. In de appendix is een voorbeeld gegeven van de berekening, uitgevoerd voor één van de strategieën.

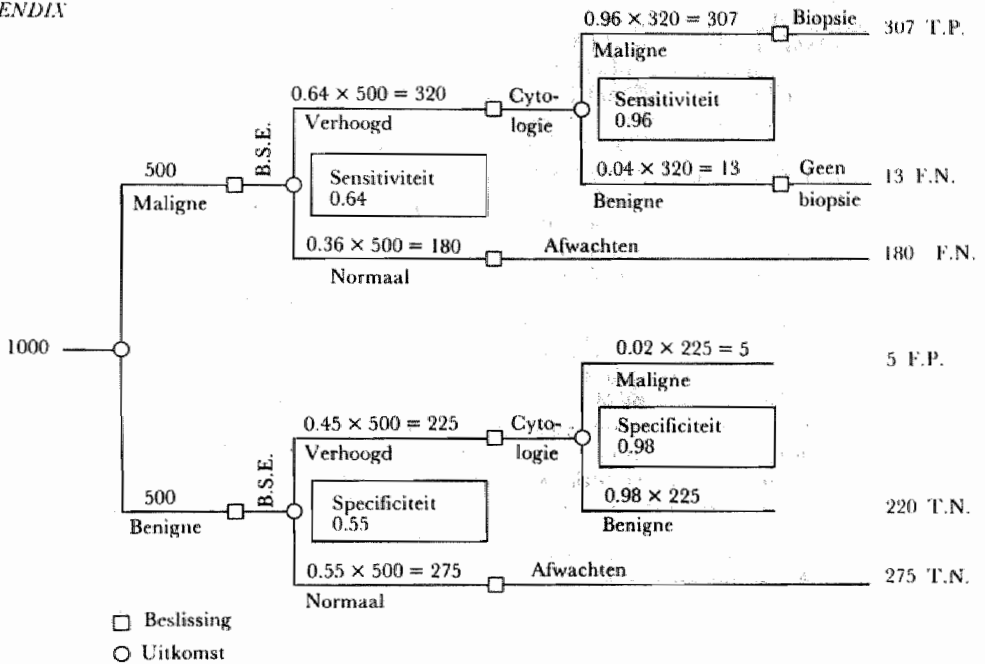
### Resultaten

Tabel 2 bevat de resultaten van de analyse. Voor iedere diagnostische strategie (1 t/m 4.2b) kan men voor huisarts en specialist, c.q. in situaties met verschillende prevalenties (a priori-kansen), afzonderlijk een indruk krijgen van de consequenties die te verwachten zijn op grond van de door ons gebruikte gegevens. Het lijkt erop dat de strategieën 2a, 3.1, 3.2a, 3.2b en 4.2a in situaties waarbij een zeer hoge prevalentie geldt (specialist) tot een meer of minder groot aantal in eerste instantie gemiste maligne afwijkingen leiden. In de situatie met de lage prevalentie (huisarts) geldt dit voor de strategieën 3.1, 3.2a en 3.2b. Hierbij gaat het om de vraag, of men de door een bepaalde strategie uitgespaarde diagnostiek vindt opwegen tegen de kans van het missen van een maligne ziekte.

Voorselectie door een B.S.E.-bepa-



APPENDIX



Voorbeeld berekening voor strategie. 3.2a bij a priori-kans van 0,50 toegepast op cohort van 1000 personen.

ling leidt telkens tot veel fout-negatieve en ook fout-positieve uitslagen en is dus geen goede strategie.

Cytologie als middel tot voorselectie om daarmee een aantal biopsieën te omzeilen, is in de situatie met de hoge prevalentie (de specialist) naar ons oordeel alleen zinvol als men de negatieve punctieresultaten niet vertrouwt en in die gevallen nog een bi-

opsie doet (2b, resp. 4.2b). Als men daarnaast toch ook nog een biopsie zou moeten doen in een groot deel van de gevallen met een positief punctieresultaat (ten behoeve van nadere classificatie), valt het voordeel weg dat veel biopsieën worden omzeild.

Cytologie, toegepast door of op aanvraag van de huisarts, is juist niet



zinnig als een negatieve uitslag gevolgd zou moeten worden door een biopsie (2b, 4.2b). Het aantal biopsies (c.q. verwijzingen) zou er niet of nauwelijks door verminderen, terwijl bovendien tevoren een groot aantal puncties zou moeten worden uitgevoerd. Cytologie zou wel van betekenis kunnen zijn, als alleen een afwijkende uitslag tot biopsie zou leiden (2b, 4.2a). Een zeer groot aantal biopsies zou worden 'uitgespaard'. Het is daarbij echter niet te voorkomen dat van de maligne aandoeningen in eerste instantie 4% wordt gemist.

Enige weken afwachten, of eventueel eerst andere gerichte diagnostiek doen (en dan vooral serologie), lijkt zinvol. Als men ervan uitgaat, en dat is redelijk, dat maligne afwijkingen daarbij in de verdachte categorie blijven, wordt hun prevalentie in de nader onderzochte groep duidelijk hoger. In de situatie met de hoge prevalentie (specialist) lijken vooral de strategieën 4.1 en 4.2b nuttig. Daarbij geldt weer dat dit nut van 4.2b t.o.v. 4.1 afneemt naarmate het aantal biopsies ook bij positieve cytologie-uitslag toeneemt. Bij lage prevalentie (de huisarts) is 4.2a een goede optie.

### Conclusie

Concluderend kan gezegd worden, dat in situaties met een hoge prevalentie (a priori-kans), zoals we die zien in klinisch-specialistische series, de tussenstap van punctie-cytologie zinvol kan zijn, indien een negatieve

cytologie gevolgd wordt door biopsie. Als voorwaarde geldt dan, dat positieve cytologie in een groot aantal van de gevallen biopsie overbodig maakt. Het is een kwestie van beoordeling hoeveel uitgespaarde biopsies men vindt opwegen tegen het altijd verrichten van cytologische punctie. Het rendement van deze strategie kan verhoogd worden als men haar toepast op die gevallen waarbij, na een periode van afwachten eventueel aangevuld met andere diagnostiek, geen genezing is opgetreden of geen afdoende verklaring gevonden is.

Wat de huisarts betreft, bij wie een lage a priori-kans te verwachten is, lijkt cytologie zinvol als een negatieve uitslag daarvan in principe *niet* tot biopsie leidt. Niet te voorkomen is dan, dat in eerste instantie een gering aantal maligne afwijkingen wordt gemist. Men dient weer te beoordelen, of dit gecompenseerd wordt door een groot aantal uitgespaarde biopsies. Ook hier geldt, dat deze strategie efficiënter wordt als men haar concentreert op een groep patiënten die geselecteerd is op grond van een voorafgaande periode van *afwachten* c.q. eenvoudige diagnostiek. Een bloedbezinking als methode van voorselectie voor cytologie of punctie levert naar verwachting te veel fout-negatieven op.

### Slotopmerkingen

Situaties kunnen verschillen, evenals a priori-kansen. Men kan de analyse daarop afstemmen. Zo kan men de a priori-kans variëren al naargelang de

leeftijd van de patiënt of de lokalisatie van de klier. Ook kan men zijn veronderstellingen anders stellen. Er is bij voorbeeld een publicatie die tot een lagere sensitiviteit (89%) en een nog hogere specificiteit (100%) voor cytologie komt dan wij hanteerden.<sup>12</sup> Bij doorrekening van deze cijfers in onze analyse wijzigt het beeld in principe niet. Bedacht moet worden dat het verrichten van cytologische punctie vaardigheid en ervaring vereist. De sensitiviteit zal lager worden als men deze ingreep, bij voorbeeld als huisarts, te weinig verricht. In sommige centra is het mogelijk om de punctie als speciale diagnostische verrichting aan te vragen bij de patholoog-anatoom.

Ook kunnen in principe de therapeutische consequenties (verbetering van prognose; mate van belasting door de therapie; bijwerkingen; ten onrechte wel c.q. niet behandeld worden; de consequenties van enkele weken wachten) in de analyse worden betrokken.

Om tot een exacte afweging te komen, kan aan iedere einduitkomst én aan de weg waarlangs deze bereikt is (zoals de hoeveelheid verrichte diagnostiek; belang van 'delay') een waarde worden toegekend. Per strategie kan dan een verwachte waarde worden berekend. De verhouding tussen de verwachte waarden van de diverse strategieën kan op haar stabiliteit getest worden door de veronderstellingen over een redelijk traject te variëren (sensitiviteitsanalyse).

De hier gepresenteerde klinische

uitkomsten kunnen door de lezer zelf van waarden worden voorzien en eventueel worden herberekend voor verschillende situaties. De teneur van de conclusies – waarbij vooral de consequenties van het verschil tussen de patiëntenpopulaties van huisarts en specialist voorop staan – zal echter duidelijk zijn. □

## Literatuur

1. Allhiser, J.N.; McKnight, T.A.; Shank, J.C.: *Lymphadenopathy in a Family Practice*. The J. of Fam. Practice, 1981; 12: 27-32.
2. Greenfield, S.G.; Jordan M.C.: *The clinical investigation of lymphadenopathy in Primary Care Practice*. J.A.M.A., 1978; 240: 1398-1399.
3. Anthony, P.P.; Knowless, S.A.S.: *Lymphadenopathy as a primary presenting sign: a clinicopathological study of 228 cases*. Br. J. Surg., 1983; 70: 112-114.
4. Lamberts, H.: *Incidentie en prevalentie van gezondheidsproblemen in de huisartspraktijk*. Huisarts en Wetenschap, 1982; 25: 401-414.
5. Saltzstein, S.L.: *The fate of the patients with non diagnostic lymph node biopsies*. Surgery, 1965; 58: 659-662.
6. Lake, A.M.; Oski, F.A.: *Peripheral lymphadenopathy in childhood*. Am. J. Dis. Child., 1978; 132: 357-359.
7. Sinclair, S.; Beckman, E.; Ellmann, L.: *Biopsy of enlarged superficial lymph nodes*. J.A.M.A., 1974; 228: 602-603.
8. Kassam, J.M.; Gephardt, G.N.: *Lymphadenopathy in childhood: long term follow-up in patients with non-diagnostic lymph node biopsies*. Hum. Pathol., 1974; 5: 431-439.
9. Schultenover, S.J. et al.: *Needle aspiration biopsy: role and limitations in surgical decision making*. Am. J. Clin. Path., 1984; 82: 105-110.
10. Knottnerus, J.A.; Kerkhof, P.D.; Schouten, L.J.: *De hoge bezinking*. Huisarts en Wetenschap, 1981; 27: 123-129.
11. Betsill, W.L.; Hajdu, S.I.: *Percutaneous aspiration biopsy of lymph nodes*. Am. J. Clin. Pathol., 1980; 471-479.

# Diagnostische hypothesen en a priori kansen in de huisartsgeneeskunde

J. A. KNOTTNERUS\*

In hoeverre zijn bepaalde klachten en symptomen indicatief voor bepaalde aandoeningen? Het antwoord op die vraag is van belang voor de dagelijkse praktijk van de huisarts; het is basisinformatie die nodig is voor het stellen van de indicatie tot nader onderzoek en voor het interpreteren van de uitkomsten van dat onderzoek. Meer algemeen gesproken is die informatie gewenst voor de onderbouwing van diagnostische protocollen en strategieën. In veruit de meeste gevallen ontbreekt de gewenste informatie echter en vaak werkt de huisarts met schattingen die zijn gebaseerd op een specialistische situatie. De bestaande huisartsregistratiesystemen kunnen echter worden gebruikt als gegevensbronnen voor meer op de huisartspraktijk toegesneden schattingen.

## Inleiding

Wat is de kans dat een 55-jarige rokende man met prikkelhoest een longcarcinoom heeft, en hoe ziet de differentiële diagnose van de huisarts eruit als zijn patiënt twee weken hees is? Is bij bloedverlies per anum altijd rectosigmoidoscopia en een colonfoto aangewezen. Hoe sterk is eigenlijk het verband tussen bepaalde mictieklachten en prostaathypertrofie?

In hoeverre deze - en honderden andere - klachten en symptomen indicatief zijn voor bepaalde aandoeningen, is van belang voor de dagelijkse praktijk van de huisarts. Het is basisinformatie die nodig is voor het stellen van de indicatie tot nader onderzoek en voor het interpreteren van de resultaten van dat onderzoek. Meer algemeen gesproken is deze informatie gewenst voor de onderbouwing van diagnostische protocollen en strategieën.

De meeste van deze vragen zijn op dit moment echter niet te beantwoorden en de huisarts moet daarom met schattingen werken. Vaak is hij zich er niet eens van bewust dat hij een schatting maakt en vaak berusten zijn schattingen op de

- geselecteerde - ervaring van specialisten. Het is dus zaak een beter inzicht te ontwikkelen in de kansen waarmee de huisarts te maken krijgt.

Dit artikel levert daartoe een aantal bouwstenen.

## A priori kansen

De *a priori* kans op een bepaalde aandoening is de kans op die aandoening, voordat relevante diagnostische informatie bekend is (bijvoorbeeld het antwoord op een anamnesevraag of de uitslag van een diagnostische test). Anders geformuleerd: de *a priori* kans op een aandoening is het percentage personen dat de aandoening heeft in een (denkbeeldige) populatie van personen met dezelfde kenmerken en klachten als de betreffende patiënt.

Het schatten van deze *a priori* kans is één van de elementen voor het vormen van diagnostische hypothesen.

In het algemeen kan men de patiëntenpopulaties van respectievelijk huisarts en specialist omschrijven als hun 'contactpopulaties': populaties van personen die de huisarts of, na verwijzing, de specialist consulteren met een vraag of klacht. Hierbij kunnen verschillende patronen voorkomen, die slechts gedeeltelijk afhankelijk zijn van de objectieve gezondheidstoestand: mensen die vaak komen, mensen die met verschillende klachten tegelijk komen, mensen die (al dan niet gericht) voor controle komen en mensen die nooit komen. Zo bezoekt in de loop van één jaar tijd circa 30 procent van de ingeschreven patiënten de huisarts niet, en 30 procent van de patiënten neemt 80 procent van de contacten voor zijn rekening.<sup>1</sup>

De 'contactpopulatie' betreft het niveau van de meest directe betrokkenheid van de arts, maar er zijn meer relevante niveaus te onderscheiden:

1. De algemene bevolking. Voor de huisarts is dat de gehele bij hem ingeschreven patiëntenpopulatie.
2. De deelpopulatie van personen die een bepaalde klacht hebben.

3. De deelpopulatie van personen die op de huisarts een beroep doen: de huisartscontactpopulatie.

4. De deelpopulatie van personen die de klacht presenteren bij de huisarts. De *a priori* kans in deze deelpopulatie, valt samen met de voorspellende waarde van de klacht zoals gepresenteerd bij de huisarts, mede in het licht van de voorkennis van de huisarts over de patiënt.

5. De deelpopulatie van personen die naar aanleiding van de klacht op de ziekte worden onderzocht door de huisarts.

6. De 'contactpopulatie' van de specialist, met weer een eigen *a priori* kans. Deze *a priori* kans is gelijk aan de *a posteriori* kans die gepaard gaat aan het besluit van de huisarts om te verwijzen.

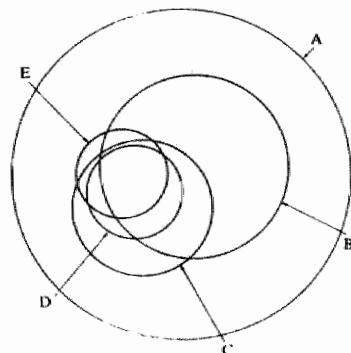
7. De deelpopulatie van personen die door de specialist onderzocht worden vanwege de klacht.

8. De deelpopulatie van patiënten die in het ziekenhuis wordt opgenomen.

Steeds gaat het om populaties met gemeenschappelijke (diagnostische) kenmerken en een eigen *a priori* kans. Deze deelpopulaties kunnen elkaar gedeeltelijk overlappen (2 en 3) of een deelverzameling van elkaar zijn (4 van zowel 3 als 2; 6 van 3, 8 van 6, 7 van 6, 7 van 5, 5 van 4) (figuur).

Tussen de deelpopulaties kunnen belangrijke verschillen bestaan wat betreft de *a priori* kans op ziekte, en bij de overgang van minder naar meer geselecteerde deelpopulaties kunnen allerlei

**Figuur.** Het selectieproces van de algemene praktijkbevolking naar de contactpopulatie van de specialist.



### Toelichting

A - praktijkbevolking; B - klacht; C - bezoekt huisarts (contactpopulatie huisarts); D - verwezen naar specialist (contactpopulatie specialist); E - ziek. De *a priori* kans op ziekte in een bepaalde deelpopulatie is steeds de proportie van de met die deelpopulatie corresponderende cirkel, die door cirkel E (ziekte) wordt bezet.

\* Arts-epidemioloog, capaciteitsgroep Huisartsgeneeskunde, Rijksuniversiteit Limburg.

veranderingen optreden in de aan te treffen verbanden tussen enerzijds klachten (symptomen) of testuitslagen en anderzijds ziekte.<sup>2</sup> Bovendien zullen de ernst en de prognose van aandoeningen samenhangen met de mate van selectie. Zo blijkt dat de kans op latere recidief-convulsies na koortsconvulsies veel kleiner is als men deze baseert op onderzoeken in de algemene bevolking, dan als men uitgaat van een populatie kinderen die bij een specialist of in ziekenhuis onder behandeling is.<sup>3</sup> Elk niveau van gezondheidszorg krijgt te maken met slechts het topje van de ijsberg van het vorige echelon: slechts ongeveer 10 procent van de gezondheidsklachten leidt tot doktersbezoek,<sup>4,5</sup> en van de ziekte-episoden waarmee de huisarts te maken krijgt, wordt vervolgens maar ongeveer 15 procent verwezen<sup>6</sup> (bijlage).

In het algemeen geldt dat de *a priori* kans op ziekte groter wordt naarmate de selectie naar hogere echelons voortschrijdt. Dit geldt met name voor die aandoeningen waarvoor specialistische behandeling wenselijk wordt geacht. Zo zien we in de tabel hoe sterk de *a priori* kansen op diverse toestanden kunnen verschillen, als men uitgaat van acute buikpijn en de cijfers vergelijkt voor presentatie in respectievelijk het eerste echelon, op de eerste hulp en op een chirurgische afdeling.<sup>7</sup>

### Multipale a priori kansen

Meestal is er niet slechts één diagnostische hypothese aan de orde, maar worden verschillende aandoeningen overwogen.<sup>8</sup> Deze kunnen afzonderlijk, maar ook in combinatie bij één patiënt voorkomen. Een van de hypothesen kan voorts zijn: het niet voorkomen van ziekte. In schema kunnen deze hypothesen, met hun respectieve *a priori* kansen, als volgt worden weergegeven:

Hypothese	<i>A priori</i> kans
D <sub>1</sub>	P(D <sub>1</sub> )
D <sub>2</sub>	P(D <sub>2</sub> )
D <sub>3</sub>	P(D <sub>3</sub> )
⋮	⋮
D <sub>i</sub>	P(D <sub>i</sub> )
⋮	⋮
D <sub>1</sub> , D <sub>i</sub>	P(D <sub>1</sub> , D <sub>i</sub> )
⋮	⋮
D <sub>1</sub> , D <sub>j</sub> , ..., D <sub>k</sub>	P(D <sub>1</sub> , D <sub>j</sub> , ..., D <sub>k</sub> )
⋮	⋮
Gezond (D <sub>-</sub> )	P(D <sub>-</sub> )

Dit schema is in feite een weergave van de differentiële diagnose, met de aantekening dat bij de differentiële diagnose

**Tabel. Verschillen in het patroon van aandoeningen bij patiënten met acute buikpijn, die zich presenteren op respectievelijk een chirurgische afdeling, de eerste hulp en bij de huisarts. Percentages.**

Diagnose	Presentatie bij		
	chirurgische afdeling	eerste hulp	huisarts
Appendicitis	28	16	3
Diverticulitis	1	1	
Perforatie ulcus pepticum	2	2	
Cholecystitis	9	5	
Obstructie	4	3	
Pancreatitis	2	1	1
Urologische afwijking	—	6	
Gynaecologische afwijking	—	5	
Aspecifieke klachten	55	62	95

Bron: De Dombal.<sup>9</sup>

niet alleen een kans-aspect, maar ook nog een waarde-aspect van de betreffende diagnoses aan de orde kan zijn.

De som van deze *a priori* kansen is 100 procent, indien alle relevante (combinaties van de diverse) hypothesen worden opgesomd. Indien men dit niet doet kan de som hoger dan 100 procent uitkomen, zoals blijkt uit het volgende voorbeeld:

Een 54-jarige man bezoekt de huisarts wegens gebrek aan eetlust sinds ongeveer vier maanden, en enig gewichtsverlies. Kort na het eten heeft hij pijn in epigastrio en een branderig gevoel achter het borstbeen. Als hij gaat liggen wordt dit branderige gevoel erger. Hij heeft het erg druk gehad het afgelopen jaar want hij is benoemd tot chef van het warenhuis waar hij al jaren werkt.

Stel dat voor deze man nu de volgende hypothesen worden overwogen:

- geen aandoening (D<sub>-</sub>);
- ulcus ventriculi (UV);
- reflux oesphagitis (RO);
- carcinoma ventriculi (CV).

Stel dat verder de volgende *a priori* kansen gelden:

P(D <sub>-</sub> ) =	45 %
P(UV + RO) =	15 %
P(UV alleen) =	10 %
P(RO alleen) =	10 %
P(UV + CV) =	10 %
P(UV + CV + RO) =	5 %
P(CV alleen) =	5 %

Het totaal bedraagt dan 100 %.

Als men geen rekening houdt met combinaties, krijgt men:

P(D <sub>-</sub> ) =	45 %
P(UV) =	40 %
P(RO) =	30 %
P(CV) =	20 %

De 'som' zou nu zijn: 135 %.

Er moet dus rekening worden gehouden met meer of minder waarschijnlijke combinaties van aandoeningen en het is daarbij wenselijk elke combinatie van een eigen kansschatting te voorzien. In het voorbeeld zien we ook een demonstratie van het niet van elkaar onafhankelijk zijn van aandoeningen (en dus van *a priori* kansen): de kans dat UV en CV gezamenlijk voorkomen, is 15 procent, terwijl dit bij 'onafhankelijkheid' slechts 8 procent zou zijn (40 procent van 20 procent).

Deze uitbreiding over 'multipale' *a priori* kansen is van belang, omdat zij aangeeft dat het werken met één *a priori* kans op één enkelvoudige ziekte-toestand - zoals in de medische beslissende vaak gebeurt - in feite te eenvoudig is. Een klachtenpatroon of een klinisch beeld laat in het algemeen meer dan één diagnostische hypothese toe, en hierop dient men ook terug te komen bij het bepalen van de *a posteriori* kans (predictieve waarde) op grond van een testuitslag. Men kan dan voor iedere hypothese apart de *a posteriori* kans bepalen.<sup>9,10</sup> Met name in de huisartsensituatie is de probleemruimte nog erg groot en dienen uiteenlopende hypothesen te worden overwogen.

## Verwachte kansen

Als artsen diagnostische hypothesen vormen, weten ze niet alleen (nog) niet wat er met de patiënt aan de hand is, ze kennen in het algemeen ook de *kans* op diverse aandoeningen niet of niet precies. Er is immers onvoldoende onderzoek gedaan om deze kansen te kunnen bepalen van iedere patiënt met diens klinisch beeld (v.t. van iedere denkbare of hypothetische populatie).

Toch maken artsen zich een voorstelling van de *a priori* kans, en zij kunnen dit alleen doen door deze te schatten. Naast  $P(D)$ , de *a priori* kans op ziekte  $D$ , moeten we dus onderscheiden:  $E[P(D)]$ : de verwachte (expected) *a priori* kans.

Ook hier geldt dat verschillende diagnostische hypothesen kunnen worden geformuleerd, die afzonderlijk of gecombineerd kunnen voorkomen, met hun respectieve geschatte/verwachte kansen. Idealiter zal  $E[P(D)]$  gelijk zijn aan  $P(D)$ , maar in de praktijk is dit (nog) lang niet altijd haalbaar of controleerbaar.

## Werkhypothesen

Er is een belangrijk verschil tussen waarschijnlijkheidsdiagnosen en werkhypothesen. In waarschijnlijkheidsdiagnosen komt alleen het kansaspect aan de orde. In de formulering van werkhypothesen wordt ook een waarde-aspect verwerkt: de consequenties van het al dan niet stellen van een bepaalde diagnose. Het zal daarbij vooralsnog om: – de ernst van de overwogen diagnose(n); – de behandelbaarheid.

Ernstige aandoeningen zal men niet graag willen missen, zeker niet als ze goed te behandelen zijn. Als ze niet zijn te behandelen, zijn er vaak andere (bijvoorbeeld sociale) redenen waarom het van belang is om ze vast te stellen.

Bij een rokende man van 55 jaar die een kriebelhoest heeft, zal men een ongecompliceerde tracheïtis veel waarschijnlijker achten dan een longcarcinoom. Tracheïtis is dus de waarschijnlijkheidsdiagnose. Toch zullen veel artsen een thorax-foto maken om longcarcinoom uit te sluiten. 'Longcarcinoom' is dan de werkhypothese.

De werkhypothese is dus de uitdrukking van een 'kansaspect', en van een 'waarde-aspect' (de consequenties betreffend) of meer precies: een subjectief geschatte kans  $E[P(D)]$  en een subjectief geschatte waarde  $E(W)$ . Deze laatste is sterk gekleurd door de individuele

situatie, omdat de afweging omtrent de ernst van het beeld, het belang van nadere diagnostiek en het nut van behandeling voor iedere patiënt anders kan uitvallen. De arts maakt ook hiervan telkens een beoordeling, in overleg met de patiënt. Het gaat ook hier weer om psychologische dataverwerkingsprocessen, die nadere bestudering vergen.

Het ligt voor de hand dat de vorming van werkhypothesen afhankelijk is van de setting waarbinnen men werkt. In hun proefschrift doen Gerritsma en Smal verslag van hun onderzoek naar verschillen in werkwijze tussen huisartsen en internisten.<sup>11</sup> Wat betreft de vorming van 'werkhypothesen' vonden zij dat internisten vaker werkhypothesen formuleren over ziektebeelden die in de huisartspraktijk weinig voorkomen. Bovendien concludeerden zij dat huisartsen zich in hun onderzoek voornamelijk door werkhypothesen laten leiden, terwijl internisten frequenter gebruik maken van routineprocedures. Zowel huisartsen als internisten gaven spontaan (zonder geheugensteun) aan – gemiddeld – zes à zeven werkhypothesen tijdens het consult gehanteerd te hebben.

Een kanttekening is, dat Gerritsma en Smal spreken over 'werkhypothesen' zonder hierbij onderscheid te maken tussen het 'kansaspect' en 'waarde-aspect'. Onlangs konden wij echter overeenkomstig bevindingen melden met gebruik van een ordening van diagnostische hypothesen alleen naar waarschijnlijkheid.<sup>12</sup>

## Bronnen voor de schatting van $P(D)$

$P(D)$  is meestal onbekend, zeker bij subpopulaties met specifieke kenmerken en over de *a priori* kans op in combinatie voorkomende aandoeningen is nog minder bekend. En zelfs al is  $P(D)$  in principe bekend, dan is het nog de vraag in hoeverre individuele medici deze kennis op het juiste moment paraat hebben of snel genoeg kunnen mobiliseren.

Kennis van  $P(D)$  – en het bevorderen dat  $E[P(D)]$  daarop aansluit – is van wezenlijk belang voor de functie van de huisarts als 'poortwachter' van de gezondheidszorg. Bovendien is het kennen van de kansen van belang, omdat ze uitgangspunt zijn van beslissende analyses, bijvoorbeeld ten behoeve van protocolontwikkeling. Daarbij kan het immers nodig zijn om op grond van de *a*

*priori* kans en de sensitiviteit en de specificiteit van een test voorspellende waarden te bepalen.<sup>13</sup>

Schattingen van  $P(D)$  in de huisartspraktijk kunnen worden gebaseerd op: 1. Registraties van klachten, bijvoorbeeld met de Reason For Encounter Classification (RFE-C) of de International Classification of Primary Care (ICPC).<sup>14-16</sup>

2. Registraties van het morbiditeitsaanbod, zoals het Monitoringproject<sup>17</sup> en de Nijmeegse Continue Morbiditeits Registratie.<sup>18</sup> Door toepassing van de ICPC (componenten 1 en 7) of een combinatie van RFE-C en ICHPPC-2<sup>19</sup> kan verder informatie worden verkregen over zowel klachten als gediagnosticeerde aandoeningen.<sup>14-20</sup>

3. Registratieprojecten gericht op een beperkt aantal aandoeningen. Hiertoe behoren de peilstationprojecten van NIVEL, Amsterdam, Den Haag en Rotterdam.<sup>21-25</sup>

4. Gericht onderzoek aangaande een bepaalde aandoening. Voorbeelden hiervan zijn het IMIR-project (Imminent Myocardial Infarction Rotterdam),<sup>26</sup> de projecten van de Dombal over de diagnostische betekenis van de 'acute buik',<sup>27-28</sup> een eigen onderzoek naar het verband tussen onverklaarde moeheidsklachten en het Hb-gehalte,<sup>29</sup> en het in Maastricht lopende onderzoek naar oorzaken van acute diarree in de huisartspraktijk.<sup>30</sup>

De meeste morbiditeitsregistraties zoals bedoeld onder 2 en 3 leveren incidenties en prevalenties op in de totale basispopulatie, inclusief de gezonden en de niet-sprekbezoekers. In dat geval zijn ze minder geschikt voor schattingen van *a priori* kansen in de contactpopulatie van de huisarts; en dat geldt zeker voor schattingen van kansen op basis van bepaalde klachten en symptomen ( $P(D|S)$ : de kans op  $D$ , gegeven een bepaalde symptomatologie  $S$ ).

Gecombineerde registratie van gepresenteerde klacht(en) en eventueel gestelde diagnosen kan echter wel de in de praktijk gewenste *a priori* kans opleveren. Genoemd is reeds de koppeling van de RFE-C aan de ICHPPC-2. Op dit moment wordt in Maastricht het Registratienet Huisartsenpraktijken ontwikkeld. Door middel van gelijktijdige registratie van subjectieve gegevens en diagnose (met diagnostische criteria) zal ook hier een schatting van  $P(D|S)$  bereikbaar zijn.<sup>30</sup> Ook onderzoek zoals onder 4 vermeld kan een schatting van  $P(D|S)$  opleveren.

## Beperkingen

Bij het gebruiken van registratiesystemen als gegevensbron voor de kansen P(D) en P(D|S) bij mensen die de huisarts bezoeken, doen zich de volgende problemen voor:

- Sluitende, in de tijd consistente en generaliseerbare afspraken over de criteria die aangehouden dienen te worden voor het registreren van bepaalde klachten en aandoeningen, zijn niet altijd gemaakt dan wel in de praktijk – op controleerbare wijze – na te leven.
- Routinematige – dat wil zeggen vooraf niet op een bepaalde vraagstelling gerichte – registratie is gebonden aan zodanige praktische beperkingen (uitgebreidheid, bewerkelijkheid, tijd), dat voor de medicus practicus relevante detailleringen van klachten niet of nauwelijks kunnen worden aangebracht. Zo wordt in de ICPC wel gesproken over 'pain attributed to heart', maar of deze uitstraalt, en zo ja, bij welke gelegenheid en hoe lang ze optreedt, of ze reageert op nitrobaat, dat is niet aan te geven in het kader van dat systeem.
- Het is vaak niet te vermijden dat een probleem onder meer dan één code geregistreerd zou kunnen worden, en dat een subjectieve keuze gemaakt wordt door de arts. Zo is duidelijk dat men sommige klachten van vrouwen zowel onder 'nervus-functioneel' als onder 'climacterieel' zou kunnen onderbrengen. Dit kan leiden tot onderschatting van het voorkomen van een bepaalde klacht of klachtengroep.
- Daarnaast komt het voor dat bepaalde klachten die men geïsoleerd zou willen beschouwen, tezamen met andere onder één code worden geplaatst. Zo worden in de ICPC menorrhagie en polymenorrhoe, toch duidelijk verschillende entiteiten, beide onder code X03 geregistreerd.
- Registratie geschiedt vaak door de onderzoekende (en waarnemende) arts zelf, en we kunnen dan, mede gezien de vaak slechts globaal te maken afspraken, te maken krijgen met een aanzienlijke inter- en intra-observervariatie: als men dezelfde patiënt door meer artsen zou laten 'scoren' zouden verschillende coderingen kunnen worden gekozen. Een tweede probleem is de detectiebias: men stelt al dan niet een indicatie voor diagnostisch onderzoek en sommige artsen zullen een bepaalde aandoening eerder opsporen en dus 'scoren' dan andere. Bovendien kan (de kennis over) het wel respectievelijk niet aanwezig zijn van een bepaalde klacht

leiden tot een verschillende intensiteit van de diagnostiek gericht op bepaalde aandoeningen, zodat de resulterende informatie over het verband tussen klacht en diagnose vertekend kan zijn. Deze detectiebias zal sterker dreigen, naarmate de kans op geheel of gedeeltelijk subklinisch verloop van een ziekte groter is. Een voorbeeld kan zijn de opsporing van ritmestoornissen.

## Uitdagingen

De genoemde problemen vormen evenzovele uitdagingen aan de onderzoekers die bij registratiesystemen betrokken zijn, om hiervoor oplossingen te zoeken. Zo wordt er op dit moment op verschillende plaatsen in ons land aan gewerkt om diagnostische criteria af te spreken en toe te passen. Speciale studies zijn echter nodig wanneer het om klachten gaat met veel relevante subdetailleringen (naar aard, lokalisatie, begeleidende verschijnselen, voorgeschiedenis en beloop) en om diagnosen die niet zonder meer op grond van eenmalig af te spreken, eenduidige en door de huisarts eenvoudig toe te passen criteria te stellen zijn. Dit zal bijvoorbeeld gelden voor beantwoording van de vraag in hoeverre bepaalde pijnklachten indicatief zijn voor het bestaan van osteoporose. Zulke speciale studies dienen te bevatten:

- gedetailleerde afspraken met betrekking tot de definiëring en precisering van de klacht(en);
- een gestandaardiseerd diagnostisch en/of follow-up protocol ter opsporing van de eventueel aanwezige relevante aandoeningen.

De IMIR-studie, gericht op de diagnostiek van het hartinfarct, is een goed voorbeeld van een onderzoek waarin aan deze vereisten veel aandacht is besteed.<sup>26</sup>

Het nadeel dat ten aanzien van een onderzoek gericht op één of enkele klachten en/of diagnosen dikwijls wordt genoemd, is dat bij een 'beperkte registratie' de arts te weinig registratieroutine zou krijgen. Hierdoor zou de arts te midden van de dagelijkse drukte kunnen vergeten patiënten te selecteren voor de studie. Anderzijds zou door concentratie op een beperkt aantal items ook te veel aandacht naar deze items kunnen uitgaan, met als gevolg overdetectie en overschatting van hun voorkomen.

Juist bij een beperkte opzet kunnen deze bezwaren worden ondervangen door gerichte afspraken en een goede controle hierop.

Welke aanpak op welk moment het meest geschikt en het meest efficiënt is, hangt af van het onderwerp en de praktische mogelijkheden. Registratie kan afdoende zijn als het gaat om vooraf goed gedefiniëerde klachten en aandoeningen die vrijwel altijd in korte tijd manifest worden en onduidelijk zijn vast te stellen. Meer gericht onderzoek is in elk geval wenselijk als het klinisch van belang is meer te weten te komen over de kans op aanwezigheid van ziekten, die meer gevarieerd en over een langere periode verlopen en waarvan de diagnostiek gepaard gaat met sterke observervariatie.

<sup>1</sup> Van Es JC. Patiënt en huisarts. Een leerboek huisartsgeneeskunde. Utrecht: Oosthoek, Scheltema & Holkema, 1974.

<sup>2</sup> Knottnerus JA, Knipschild PG, Sturmans F. Symptomatie en selectiebias. Huisarts en Wetenschap 1985; 28: 325-30.

<sup>3</sup> Ellenberg JH, Nelson KB. Sample selection and the natural history of diseases: studies of febrile seizures. JAMA 1980; 243: 1337-40.

<sup>4</sup> Folmer HR. Huisarts en ijsberg [Dissertatie]. Utrecht: Rijksuniversiteit Utrecht, 1968.

<sup>5</sup> Huygen FJA, Van der Hoogen H, Neefs WJ. Gezondheid en ziekte; een onderzoek van gezinnen. Ned Tijdschr Geneesk 1983; 127: 1612-9.

<sup>6</sup> Van Es JC, Pijlman HR. Het verwijzen van ziekenfondspatiënten in 122 Nederlandse huisartspraktijken. Huisarts en Wetenschap 1970; 13: 433-49.

<sup>7</sup> De Dombal FT. Transporting databanks of medical information from one location to another. Effective Health Care 1983; 1: 155-62.

<sup>8</sup> Elstein AS, Shulman LS, Sprafka SA. Medical problem solving. An analysis of clinical reasoning. Cambridge (Ma.): Harvard University Press, 1978.

<sup>9</sup> Weinstein MC, Fineberg HV. Clinical decision analysis. Philadelphia: Saunders, 1980.

<sup>10</sup> Knottnerus JA, Sturmans F. Het belang van de niet-afwijkende bevinding voor de differentiële diagnostiek. Practitioner 1984; 1: 147-51.

<sup>11</sup> Gerritsma JGM, Smal JA. De werkwijze van huisarts en internist. Een vergelijkend onderzoek met behulp van een interactieve patiëntensimulatie. Utrecht: Bunge, 1982.

<sup>12</sup> Boenink A, De Vos P, Knottnerus A. Moehheid: verschillen tussen huisartsen en internisten. Med Contact 1986; 41: 54-7.

<sup>13</sup> Knottnerus JA. Principes van beslistkunde. Nieuw kompas voor de huisarts, juli 1984, II-4 – II-15.

<sup>14</sup> Lamberts H. Redenen om naar de huisarts te gaan. Eerste ervaringen met de Reason for Encounter Classification. Huisarts en Wetenschap 1982; 25: 301-10.

<sup>15</sup> Anonymous. Reason For Encounter Clas-

- sification. Field trial manual 1983.
- <sup>16</sup> Anonymous. International Classification of Primary Care (ICPC). Manual for use in relevance studies. Amsterdam: WHO-working party on the ICPC, 1985.
- <sup>17</sup> Lamberts H. Morbidity in general practice. Diagnosis related information from the monitoring project. Utrecht: Huisartsenpers, 1984.
- <sup>18</sup> Continue Morbiditeits Registratie Nijmegen. Werkgroep Epidemiologie in de huisartspraktijk 1971-1978. Gewone ziekten. Nijmegen: Nijmeegs Universitair Huisartsen Instituut, 1980.
- <sup>19</sup> ICHPPC-2-Defined (Inclusion criteria for the use of the rubrics of the International Classification of Health Problems in Primary Care). Oxford: Oxford University Press, 1983.
- <sup>20</sup> Van Weel C. De visie van de patiënt. Gegevens ontleend aan een profonderzoek met de RFEC. Huisarts en Wetenschap 1982; 25(suppl H&P 6): 30-4.
- <sup>21</sup> Anoniem. Continue Morbiditeits Registratie Peilstations Nederland. Jaarverslag 1983. Utrecht: Nederlands Huisartsen Instituut, 1984.
- <sup>22</sup> Collette HJA, Bijkerk H. Vijftien jaar peilstations Nederland 1970-1984. Huisarts en Wetenschap 1985; 28: 207-10.
- <sup>23</sup> Anoniem. Amsterdams Peilstation Project. Jaarverslag 1983. Amsterdam: GG & GD, 1984.
- <sup>24</sup> Huisartsregistratiesysteem Den Haag. Bof, ziekte van Osgood-Schlatter en suïdaal gedrag [Epidemiologische notities]. Huisarts en Wetenschap 1985; 28: 276-8.
- <sup>25</sup> Epidemiologisch Bulletin. Rotterdam: Gemeentelijke Geneeskundige Dienst.
- <sup>26</sup> Van der Does E, Lubsen J. Acute coronary events in general practice. The imminent myocardial infarction Rotterdam study. [Dissertatie]. Rotterdam: Erasmus Universiteit Rotterdam, 1978.
- <sup>27</sup> De Dombal FT. Acute abdominal pain. An OMGE survey. Scand J Gastroenterol 1979; 14 (suppl 56): 29-43.
- <sup>28</sup> De Dombal FT, Leaper DJ, Horrocks JC, Staniland JR, McCann AP. Human and computer-aided diagnosis of abdominal pain: further report with emphasis on performance of clinicians. Br Med J 1974; i: 376-80.
- <sup>29</sup> Knottnerus JA, Knipschild PG, Van Wersch JWJ, Sijstermans AH. Onverklaarde moeheidsklachten en hemoglobinegehalte, een onderzoek vanuit de huisartspraktijk. Ned Tijdschr Geneesk 1986; 130: 402-5.
- <sup>30</sup> Van der Horst F, Seelen A, Vissers F, Plagge H, Von Höfen R, De Geus C, Kleijnen J. Registratie in de huisartspraktijk. Over de betrouwbaarheid, praktische hanteerbaarheid en adequaatheid van het RFE-classificatiesysteem. Huisarts en Wetenschap 1985; 28: 229-34.

**Bijlage. De situatie van huisarts en medisch specialist met betrekking tot gepresenteerde problemen, ziekten en handelwijzen.**

Huisarts	Medisch specialist
<b>Klacht/probleem</b> grote probleemruimte breed scala van problemen somatische, psychische problemen vaak verschillende problemen per contact	kleine probleemruimte probleem beperkt tot één discipline, voorgeselecteerd primair somatische problemen meestal één probleem per contact
<b>Morbiditeit</b> veel selflimiting diseases veel minor ailments vroege stadia van ziekte weinig ontwikkelde atypische symptomen relatief weinig ernstige ziekten vaak niet meer dan symptoom- of klacht-diagnose	weinig selflimiting diseases weinig minor ailments latere stadia ontwikkelde, vaak typische symptomen veel ernstige ziekten meestal diagnose van ziektebeeld
<b>Werkwijze</b> vaak expectatief probleemgeoriënteerd medische gegevens beperkt beschikbaar omgaan met onzekerheid, werken met waarschijnlijkheden	vaak snelle diagnostiek diagnose-georiënteerd volledigheid van medische gegevens wordt nagestreefd weinig onzekerheid, streven naar definitief antwoord
<b>Instrumenten</b> kennis op algemeen niveau continue zorg afwachten weinig technologie veel kennis over persoon en sociale omgeving overzicht over gehele situatie van de patiënt	somatische kennis op specialistisch niveau discontinue zorg snelle diagnostiek veel technologie weinig kennis over persoon en omgeving geen overzicht over geheel



# Moeheid: verschillen in benadering tussen huisartsen en internisten

Tussen de werksituatie van huisartsen en internisten bestaan veel verschillen. Een huisarts zal vaker patiënten zien bij wie de symptomen (nog) niet duidelijk zijn ontwikkeld. Psychosociale factoren zullen vaker meespelen bij de klacht van een patiënt. De onderzoeksmogelijkheden die een huisarts heeft zullen het niet altijd mogelijk maken een specifieke diagnose te stellen en ook zal het stellen van zo'n specifieke diagnose niet altijd nodig zijn voor het instellen van een therapie. De internisten daarentegen zien patiënten meestal in een later stadium van een ziekte, waarbij de symptomen duidelijker aanwezig kunnen zijn. Hij wordt geconsulteerd vanuit de gedachte, bij patiënt en verwijzend arts, dat de klachten een organische oorzaak hebben, waarbij het er ook om gaat zeldzame ziekten zoveel mogelijk uit te sluiten. De laboratoriefaciliteiten van een ziekenhuis bieden daartoe ook veel meer de mogelijkheid.

Gerritsma en Smal veronderstelden dat er onder meer door deze verschillen in werksituatie ook verschillen in werkwijze ontstaan<sup>1</sup>. Dit zijn zij door middel van een onderzoek nagegaan. Mogelijk door bovenstaande beschrijving verklaarbaar, is hun bevinding dat internisten bij eenzelfde casus vaker dan huisartsen zeldzame ziektes onder hun werkhypothese hadden. Met betrekking tot deze zeldzame ziektebeelden zaten er bij de huisartsen vaker werkhypothese over psychosociale oorzaken. Elstein had al eerder aangetoond dat de eerste probleemformuleringen al vroeg in het contact met de patiënt ontstaan, soms zelfs al voordat de patiënt iets heeft gezegd<sup>2</sup>. Gerritsma en Smal hebben ook geanalyseerd hoe de werkhypothese samenhangt met de door artsen gevraagde informatie. Zij herkenden drie patronen: de cascade (bij de uitwerking van een werkhypothese komt een nieuwe werkhypothese op, die meteen wordt getoetst; hierdoor wordt het oude spoor verlaten), de successieve methode (een werkhypothese wordt pas uitgediept als de toetsing van de vorige is afgerond) en de screening (in eerste instantie worden volgens een vast patroon gegevens verzameld, hetgeen niet wordt beïnvloed door de werkhypothese).

In het kader van het onderwijs aan studenten van de Faculteit der Geneeskun-

Annette Boenink,  
Paul de Vos  
André Knottnerus

*Versillen in werksituatie tussen huisartsen en internisten gaan gepaard met verschillen in werkwijze; dat is uit onderzoek gebleken. Annette Boenink, Paul de Vos en André Knottnerus, student-assistenten geneeskunde respectievelijk arts, wetenschappelijk medewerker capaciteitsgroep Epidemiologie/Gezondheidszorgonderzoek aan de RU Limburg, stelden een nader onderzoek in naar de wijze waarop huisartsen en internisten omgaan met de klacht 'moeheid'. De huisartsen bleken meer aan een psychosociale oorzaak te denken, de gedachten van de internisten gingen vaker in de richting van (ernstige) somatische pathologie. Internisten zouden door de bank genomen meer onderzoek verrichten dan huisartsen. Overigens was binnen beide groepen de onderlinge overeenstemming niet erg groot.*

de van de Rijksuniversiteit Limburg hebben wij een klein onderzoek verricht naar de manier waarop huisartsen en internisten omgaan met het probleem 'moeheid': aan welke diagnose(groepen) denken zij en wat is hun beleid in eerste aanleg? Zijn hierin verschillen te ontdekken tussen huisartsen en internisten? Met dit onderzoek wilden wij kijken of de trends die Gerritsma en Smal hebben gevonden ook kunnen worden bevestigd bij een casus over de klacht 'moeheid'.

## WERKWIJZE

Wij hebben een casus met betrekking tot de klacht 'moeheid' toegestuurd aan een aantal Zuidlimburgse huisartsen, internisten en arts-assistenten Interne Geneeskunde van het ziekenhuis St. Annadal (hierna te samen aangeduid als: internisten) zonder dat zij wisten dat zij met

elkaar zouden worden vergeleken. De casus was zodanig opgesteld dat er allerlei mogelijke oorzaken aan ten grondslag zouden kunnen liggen, van psychische problematiek tot een maligniteit:

## Casus

Mevrouw Theunissen, 55 jaar, verhuisde drie jaar geleden van Roermond naar hier, omdat haar man een nieuwe baan kreeg bij een plaatselijke bank. Deze baan vraagt nogal veel van hem. Ze hebben drie kinderen, van wie de jongste nog thuis woont. Mevrouw Theunissen komt niet vaak bij de huisarts: de laatste keer was een jaar geleden voor griep. Ze zegt dat ze altijd goed gezond is geweest, maar sinds een paar maanden voelt ze zich zo verschrikkelijk moe. Het huishouden valt haar steeds zwaarder. In de loop van de middag is ze zo moe dat ze een half uurtje moet gaan zitten. Ze heeft nergens meer energie voor. Haar man vond ook dat ze zich maar eens goed moest laten nakijken, want dit was toch niet normaal. Mevrouw vertelt verder in het gesprek dat haar laatste menstruatie anderhalf jaar geleden was en dat ze de laatste paar maanden een kilo is afgevallen.

Wij hebben met betrekking tot deze casus twee open vragen gesteld:

1. Aan welke diagnose(groepen) en problemen denkt u (in volgorde van waarschijnlijkheid)?
2. Wat is in eerste aanleg uw beleid met betrekking tot de diagnostiek?

Aan de artsen werd verzocht hun antwoorden aan ons terug te sturen.

Van de 30 benaderde internisten en arts-assistenten Interne Geneeskunde gaven er 14 gehoor aan ons verzoek; van de 27 aangeschreven huisartsen reageerden er 19. De antwoorden die zij gaven zijn door ons in bepaalde categorieën ingedeeld.

## RESULTATEN

De huisartsen bleken in het algemeen wat beter overweg te kunnen met een dergelijk vaag probleem. Van de 14 internisten gaven er 2 te kennen dat zij zo weinig gegevens hadden gekregen dat ze er nauwelijks iets zinnigs over konden zeggen; 2 andere internisten hebben wel een lijst met diagnosegroepen en problemen gegeven, maar konden niets zeggen over hun beleid met betrekking tot diagnos-



tiek zonder aanvullende gegevens over anamnese en lichamelijk onderzoek.

### Uitwerking vraag 1

Op vraag 1 werden door de huisartsen gemiddeld 4,8 diagnosegroepen genoemd (standaarddeviatie 1,5) en door de internisten gemiddeld 5,6 (standaarddeviatie 1,6). Bij de uitwerking van vraag 1 (zie tabel 1) is een onderverdeling gemaakt in groepen aandoeningen; hoe deze groepen precies zijn samengesteld is te vinden in *bijlage 1*. In kolom 1 van tabel 1 is vermeld het totaal aantal keren dat een aandoening uit een bepaalde groep werd genoemd door de huisartsen respectievelijk internisten die deze vraag hebben beantwoord. Hierbij is het mogelijk dat een bepaalde arts meer aandoeningen uit dezelfde groep noemt (zo kunnen er in één categorie meer diagnoses zijn genoemd dan er arts zijn). De tweede kolom vermeldt het aantal keren dat een aandoening uit een bepaalde groep als eerste op de lijst van een arts werd genoemd door de huisartsen respectievelijk de internisten.

Kolom 3 geeft een indicatie van het relatieve gewicht dat aan de verschillende groepen van aandoeningen werd toegekend binnen de huisartsen-respectievelijk internistengroep. Deze indicatie kwam tot stand door een bepaalde diagnose(groep)/probleem 8 punten te geven wanneer deze diagnose als eerste werd genoemd, 7 punten als ze op de tweede plaats stond, enz. Vervolgens wordt per diagnosegroep het door alle huisartsen respectievelijk internisten gegeven aantal punten opgeteld en daarna berekend als percentage van het totaal aantal punten.

### Uitwerking vraag 2

Ook bij de uitwerking van vraag 2 (zie tabel 2) zijn de antwoorden in bepaalde categorieën ingedeeld (zie voor de exacte indeling *bijlage 2*).

De huisartsen deden gemiddeld onderzoek op 3,0 gebieden (standaarddeviatie 1,9), de internisten op 4,5 gebieden (standaarddeviatie 1,5). Veel van de onderzochte artsen (8 internisten en 13 huisartsen) hebben met zoveel woorden gesteld dat er eerst een goede anamnese en een goed lichamelijk onderzoek dient plaats te vinden; dit is verder niet als zodanig in de tabel opgenomen.

Een aantal artsen (8 van de 12 internisten en 5 van de 19 huisartsen) heeft op eigen

Tabel 1. Genoemde diagnoses ingedeeld in groepen vergeleken tussen huisartsen en internisten.

	totaal aantal keren genoemd		aantal keren als eerste genoemd		relatieve gewicht*	
	huisartsen (N=19)	internisten (N=12)	huisartsen (N=19)	internisten (N=12)	huisartsen (100% = 632 pin)	internisten (100% = 396 pin)
anemie	16	10	4	6	15%	21%
endocriнологisch/metabool	12	11	0	0	12%	14%
depressie	16	6	9	3	18%	11%
menopauze	4	1	2	1	5%	2%
maligniteit	9	11	0	1	8%	17%
infectie/ontsteking	5	7	0	0	5%	8%
zumenage	5	1	0	0	5%	2%
overige						
- psychosociaal	6	5	3	1	7%	9%
- somatisch	19	15	1	0	24%	17%
					100%	100%
gemiddeld aantal antwoorden per arts	4,8 (SD 1,5)	5,6 (SD 1,6)				

\* Toelichting in de tekst. SD = standaarddeviatie

initiatief een onderscheid aangebracht tussen wat zij in eerste en in tweede instantie zouden doen. Aangezien de vraagstelling uitging van het beleid in eerste aanleg, is dat wat men zei in tweede instantie te zullen doen niet meegeteld bij de uitwerking van deze vraag. Onderaan de tabel is aangegeven welk percentage van de huisartsen onderzoek wil

doen waarvoor het doorsturen van de patiënt naar een diagnostisch centrum noodzakelijk is.

### DISCUSSIE

Gezien het betrekkelijk geringe aantal artsen in ons onderzoek vonden wij het niet zinvol statistische toetsen toe te passen. Wij richten ons in de nu volgende bespreking op de algemene trends die uit het materiaal naar voren komen.

We kunnen niet uitsluiten dat de geringere respons van de internisten onze resultaten heeft beïnvloed. Gezien het feit dat men tevoren niet op de hoogte was van de voorgenomen vergelijking tussen huisartsen en internisten, is het echter niet aannemelijk dat deze invloed sterk is geweest.

De verwachte verschillen in hypothesevorming en beleid tussen huisartsen en internisten zijn inderdaad gevonden en komen globaal overeen met wat Gerritsma en Smal hebben gevonden.

Wat de hypothesevorming betreft: hoewel vrijwel alle artsen zowel aan somatische als aan psychosociale oorzaken voor de in de casus geschetste moeheid dachten, zijn er in de volgorde die zij hebben aangebracht duidelijke verschillen te zien. Internisten zetten vaker somatische ziektebeelden op de eerste plaats (anemie en maligniteit), de gedachten van de huisartsen gaan eerder in de richting van een depressief beeld. In totaal noemden 7 van de 12 internisten

Tabel 2. Aantal huisartsen respectievelijk internisten dat een bepaald onderzoek doet.

	huisartsen (N=19)	internisten (N=10)
bloed eenvoudig	18	10*
bloed uitgebreid	2	
urine	13	4
nierfunctie	4	9
leverfunctie	3	7
faeces benzidine	1	2
schilddierfunctie	3	2
gynaecologisch		
onderzoek/PAP-smear	3	0
X-thorax	4	6
ECG	2	1
elektrolyten	0	3
bloedsuiker	3	2
gemiddeld aantal onderzoeken per arts	3,0 (SD 1,9)	4,5 (SD 1,5)
'gesprokje'	6	1
doorsturen naar een diagnostisch centrum nodig	7	-

In de groep huisartsen werd verder nog één keer genoemd: X-colon, echo kleine bekken, Informeren naar alcoholgebruik. In de groep internisten werd nog één keer genoemd: echo pancreas, bijnierfunctie.

\* Het onderscheid tussen 'bloed eenvoudig' en 'bloed uitgebreid' kon alleen bij huisartsen worden aangebracht.

SD = standaarddeviatie

(58%) een somatische oorzaak, tegen 7 van de 19 huisartsen (37%). Vrijwel alle internisten noemen ergens op hun lijst een maligniteit, tegenover ongeveer de helft van de huisartsen. Bij 'depressie' liggen deze cijfers precies omgekeerd. Internisten denken ook wel bijna altijd aan een psychosociale oorzaak, maar zij zetten deze lager op hun lijst en zij doen dat in minder specifieke bewoordingen dan de huisartsen. Huisartsen denken vaker aan verschillende vormen van psychosociale problematiek, terwijl internisten deze vaker alle onder één noemer brengen. Een voor de hand liggende verklaring voor deze verschillen op het niveau van de groep is dat de prevalentie van de verschillende oorzaken van moeheid in de interne praktijk anders is dan in de huisartspraktijk.

Hoewel we inderdaad verschillen op groepsniveau hebben gevonden, is ons ook gebleken dat er binnen beide groepen, huisartsen en internisten, maar weinig overeenstemming bestaat. Die overeenstemming is namelijk kleiner dan de tabel suggereert, aangezien bijvoorbeeld in de groep 'endocrinologisch/metabool' zowel de artsen zijn opgenomen die diabetes mellitus hebben genoemd als degenen die bijvoorbeeld aan hypothyreoïde of bijnierinsufficiëntie dachten. Dit gebrek aan onderlinge overeenstemming kan moeilijk worden verklaard uit de verschillen in prevalentie van de diverse oorzaken van moeheid tussen de ene huisartspraktijk/interne praktijk en de andere. Een verklaring zou kunnen zijn dat, hoewel we naar de volgorde van waarschijnlijkheid hebben gevraagd, sommige artsen behalve de prevalentie ook de door hen ingeschatte ernst van de mogelijke aandoeningen en de noodzaak en/of het nut van behandeling ervan een rol hebben laten spelen bij het opstellen van een hypothesenlijst; daarbij zouden individuele ervaringen van artsen met bepaalde ziekten een rol kunnen spelen.

Naar aanleiding hiervan wijzen wij op het verschil tussen (ons) onderzoek waarin naar volgorde van waarschijnlijkheid wordt gevraagd en ander onderzoek (Gerritsma en Smal) waarin de werkhypothese centraal staat: bij een werkhypothese zijn ook factoren als de ernst van een aandoening betrokken.

Wat het beleid betreft valt op dat het 'beleid in eerste aanleg' vooral screenend van aard is: vrijwel alle huisartsen doen bloedonderzoek en de meesten ook urineonderzoek, de internisten doen een volledig bloedonderzoek en bepalen

meestal bovendien de nier- en leverfuncties. Na deze routineonderzoeken volgt nog een lange lijst met meer specifieke onderzoeken, die alle slechts door enkele artsen worden gedaan. Het was met onze werkwijze niet mogelijk een patroon te herkennen in de werkwijze van de artsen (screening, cascade of successieve methode).

Zowel bij de huisartsen als bij de internisten is de onderlinge overeenstemming ook hier ver te zoeken. Over de hele lijn lijken de internisten wel wat meer onderzoek te doen dan de huisartsen; dit verschil is vooral duidelijk bij het klinisch-chemisch onderzoek (nier- en leverfuncties en elektrolyten). Dit ligt ook wel voor de hand, aangezien internisten over meer faciliteiten beschikken. Bovendien nemen huisartsen en internisten verschillende posities in binnen de gezondheidszorg: de huisarts verzorgt de eerste opvang, wanneer het probleem nog erg ongedifferentieerd is; de specialist echter krijgt meer voorgeselecteerde problemen, waarbij van hem of haar een eindoordeel wordt verwacht. Het is gezien onze resultaten aannemelijk dat deze factoren ook bij de beoordeling van een zelfde casus van invloed zijn.

Tenslotte willen wij nog even stilstaan bij het feit dat tweederde van zowel de huisartsen als de internisten die wij hebben ondervraagd, heeft aangegeven eerst een goede, uitgebreide anamnese en dito lichamelijk onderzoek te willen uitvoeren. Deze - terecht - opmerking geeft de beperkingen aan van dit onderzoek, namelijk dat de artsen gedwongen werden hypothesen en beleid te formuleren op grond van vrij summier gegevens. Dit is een wat onnatuurlijke situatie, aangezien de artsen in hun praktijk uiteraard wel de mogelijkheid hebben via anamnese en lichamelijk onderzoek meer gegevens te verkrijgen. Waar dit met betrekking tot het beleid een vertekening kan inhouden, is deze manier van gegevens verzamelen waar het de eerste gedachten omtrent diagnose(groepen) betreft misschien directer van aard dan de werkwijze die Gerritsma en Smal bij hun onderzoek hebben gevolgd: zij vroegen pas na de volledige afhandeling van een gehele casus naar de werkhypothesen die de artsen vanaf het begin hadden gevormd.

Kijken we naar andere onderzoeken en andere literatuur over moeheid, dan is de door ons gevonden diversiteit in denken en voorgenomen handelen niet zo verwonderlijk. Zolang er geen andere en concretere aanwijzingen voor pathologie

zijn, lijkt moeheid een moeilijk oplosbaar probleem te zijn. In termen van het begrippenkader van het probleemoplossen kan worden gezegd dat moeheid een grote probleemruimte heeft. Het gedeelte over moeheid in Harrisons 'Principles of Internal Medicine' biedt weinig concrete handvaten en lijkt sterk gekleurd door de interesses van de auteur van dat hoofdstuk<sup>3</sup>. De Haan en Bleys merken in de inleiding van hun boek over moeheid op dat het gaat om een onmogelijk te omschrijven begrip en dat zij daardoor noodgedwongen meer zullen spreken over alles rondom het moeheidprobleem dan over het begrip 'moeheid' zelf; zij kunnen met de door hun voorgestelde benadering van de klacht alleen een aantal richtingen (in de zin van ernstig-niet ernstig of psychisch-somatisch-fysiologisch) aangeven en zo de arts enige steun bieden bij het gericht zoeken naar een precieze diagnose<sup>4</sup>.

In het eerste contact wordt volgens Katterdahl in 15%-35% van de gevallen met anamnese en lichamelijk onderzoek geen onderliggende oorzaak gevonden<sup>5</sup>. In een tweetal onderzoeken heeft men zich speciaal gericht op de groep patiënten met moeheid bij wie de oorzaak aanvankelijk onopgehelderd blijft<sup>5,6</sup>; hieruit blijkt dat dit vaak leidt tot het doen van zeer veel onderzoek waarvan de zin niet altijd aangetoond is. In ons onderzoek zijn wij alleen beschrijvend te werk gegaan: wij hebben niet de effectiviteit van het door de artsen voorgestane beleid gemeten.

Wij hopen dat dit onderzoek een bijdrage kan leveren aan de verheldering van de diverse benaderingen van het probleem 'moeheid' in de verschillende echelons van de gezondheidszorg. □

---

*Met dank aan de internisten, arts-assistenten Interne Geneeskunde van het ziekenhuis St. Annadal en de Zuidlimburgse huisartsen die aan dit onderzoek hebben meegedaan.*

#### Literatuur

1. Gerritsma JGM, Smal JA. De werkwijze van huisarts en internist. Utrecht: Bunge, 1982.
2. Elstein AS, Shulman LS, Sprafka SA. Medical problem solving. An analysis of clinical reasoning. Cambridge (Massachusetts): Harvard University Press, 1978.
3. Petersdorf (ed). Harrison's Principles of Internal Medicine. McGraw Hill, 1983.
4. Haan M de, Bleys FC. Moetheid. Utrecht: Bunge, 1980.
5. Katerndahl DA. Fatigue of uncertain etiology. Family Medicine Review 1983; 1: 26-38.
6. Morrison JD. Fatigue as a Presenting Complaint in Family Practice. Journal of Family Practice 1980; 10: 795-801.

#### Bijlage 1

*Onderverdeling in groepen van vraag 1 met tussen haakjes de aandoeningen zoals die door de artsen zijn opgeschreven en door ons in de desbetreffende groep ondergebracht*

- anemie: bloedarmoede, gastro-enteraal bloedverlies, hematologische afwijkingen
- endocrinologisch/metabool: diabetes mellitus, hyperthyreoïdie, hypothyreoïdie, bijniersufficiëntie, Addison, metabole/endocriene pathologie
- depressie: midlife-depressie, reactieve depressie, involutiedepressie, climacterium-depressie, depressie door verhuizing, postmenopauze depressie
- menopauze, met uitzondering van depressie (hormonale overgang)
- maligniteit: Ca, maligniteit (pancreas?), kanker, maligniteit tractus digestivus, maligniteit ovariae
- infectie/ontsteking: chronische infectie, chronische ontsteking, colitis ulcerosa/Crohn, chronische hepatitis, urinewegin-

fectie, infectieuze aandoening, prodromen hepatitis

- surmenage: overbelasting, overspannenheid
- overige psychosociaal: aanpassingsproblemen, psychosomatisch, alcoholisme, niet-somatisch, psychisch, gezinscrisis, functionele moetheid, functioneel syndroom
- overige somatisch: sarcoidose, auto-immuunziekte, elektrolytenstoornis, cardiaal, nierfunctiestoornis, nierinsufficiëntie, systeem-aandoening, hypertensie, statische afwijkingen, intoxicaties, neurologische afwijkingen, collageenziekte, pulmonaal proces, hypotensie, Kahler, uremie, decompensatio cordis (zie cardiaal), myoom, colonafwijkingen

#### Bijlage 2

*Onderverdeling antwoorden vraag 2*

1. bloedonderzoek
  - a. eenvoudig = Hb, BSE, diff. (door huisarts te doen)
  - b. uitgebreid = tenminste één van de volgende: ery's, Fe, reticulo's
2. urine (glucose, eiwit, red., sediment, uricult)
3. nierfunctie (creatinine ureum)
4. leverfuncties (AFP, GTP, SGOT, GT, cholesterol)
5. faeces benzidine (occult-test)
6. schildklier (T4)
7. gynaecologisch onderzoek (PAP-smear)
8. X-thorax
9. ECG
10. 'gesprek' (slaapstoornissen, gezinsgesprek, echtpaargesprek, psychisch uitdiepen)
11. elektrolyten
12. bloedsuiker

## Medisch onderwijs

## Intercollegiale vergelijking

## Oefenen met huisartsen in opleiding

Een veelbelovend toepassingsgebied van de medische beslistkunde is de intercollegiale toetsing. Echter, bij het ontbreken van criteria kan in feite niet van toetsing worden gesproken. Een goede werkwijze is dan de intercollegiale vergelijking: het uitwisselen en bespreken van elkaars inzichten, werkwijzen en beleidsopties. Als belangrijkste leermomenten zijn in dit verband te onderscheiden:

1. het verkrijgen van inzicht in verschillen van beoordeling bij medische probleemstellingen;
2. het kunnen localiseren van die verschillen binnen het geheel van het beleid;
3. het nagaan van de consequenties van de verschillen voor het verloop van de hulpverlening;
4. het identificeren van de stappen die moeten worden genomen om tot de formulering van een optimale strategie te kunnen komen.

Intercollegiale vergelijking is een belangrijk bestanddeel van het leerproces van huisartsen in opleiding. In het volgende praktijkvoorbeeld wordt aan de hand van ervaringen in een bijeenkomst van een groep Maastrichtse huisartsen in opleiding geïllustreerd dat bij zo'n uitwisseling een beslistkundig kader als kapstok kan fungeren. Het ging daarbij vooral om de eerste twee van de zojuist genoemde leermomenten.

## WERKWIJZE

Men is er vanuit gegaan dat het rendement van een leer- of uitwisselings situatie het grootst is als de deelnemers eigen casuïstiek inbrengen en zelf de leerstof (mee) ontwikkelen<sup>1,2</sup>. Naar aanleiding van een reële casus, die behalve aan de inbrenger aan geen van de andere deelnemers te voren bekend is, komen de volgende aspecten aan de orde:

1. het schatten van de waarschijnlijkheid van diverse diagnoses of diagnosecategorien (de a priori-kansen);

J. A. Knottnerus  
Dr. H. K. Muller

*Huisartsen in opleiding kunnen al in één les, aan de hand van één enkele casus, kennis maken met de belangrijkste aspecten van de medische besluitvorming, waaronder het optreden van verschillen in beoordeling van kansen en het kunnen plaatsen van verschillen in de beoordeling van een medische probleemstelling in het geheel van het beleid. De arts J. A. Knottnerus en Dr. H. K. Muller (RUL) beschrijven zo'n les en laten zien hoe intercollegiale vergelijking in de praktijk werkt.*

2. de vertaling hiervan in het 'pluis'- of 'niet-pluis'-gevoel; en
3. de structuur van het probleem (in de vorm van een beslisboom).

## CASUS

Op een morgen bezoekt een aan de arts-assistent nog niet bekende man van 64 jaar het spreekuur met de klacht dat hij de laatste tijd steeds afvalt en daar graag 'iets' aan wil hebben gedaan. Er zijn verder geen spontane klachten. Bij de vraagverheldering blijkt dat over de vermagering een duidelijke ongerustheid bestaat, zodat wordt besloten een en ander grondig te onderzoeken. Uit de gegevens op de patiëntenkaart (van de opleider) blijkt het volgende:

- Acht maanden geleden was er een spreekuurconsult, waarvan de reden niet is genoemd. Een haemocult-test was  $3 \times$  negatief, de BSE bedroeg 40 mm/h. Er was geen bloedverlies per anum geweest.
- Een maand later was er weer een spreekuurbezoek in verband met een vermagering van 8 kg in één jaar. Verder waren er geen klachten. Een routine-urineonderzoek leverde geen afwijkingen op.
- Weer twee weken later volgde een spreekuurbezoek wegens een abces aan de stuit,

dat na verwijzing poliklinisch door de chirurg werd behandeld. Er waren geen redenen om op dat moment verder onderzoek te laten doen.

Vandaag vertelt patiënt, de laatste zeven maanden meer dan 5 kg te zijn afgevallen. Bij navraag blijkt er ook wat breiige ontlasting te zijn geweest. Verder zijn er anamnestiche geen aanknopingspunten ter verklaring van deze klachten. Bij routine-lichamelijk onderzoek met speciale aandacht voor het abdomen worden evenmin aanknopingspunten voor een oorzaak van de klachten/verschijnselen gevonden.

## HYPOTHESEN EN PRIORI-KANSEN

Nadat de casus tot zover is gepresenteerd, wordt aan de huisartsen in opleiding en hun huisarts-begeleider gevraagd op te schrijven:

- welke diagnostische hypothese(n) zij op dit moment overwegen;
- hoe groot (in procenten) zij, gegeven de casus, de kans op elk van die diagnose(n) achten; en
- in hoeverre zij het gevoel hebben dat de toestand van de gepresenteerde patiënt 'pluis' of 'niet pluis' is. Dit wordt uitgedrukt in een percentage variërend van 0% (helemaal pluis) tot 100% (in het geheel niet pluis).

Vervolgens wordt uitgewisseld wat de deelnemers hebben opgeschreven. Wat de gegeven casus betreft bleek dat alle deelnemende artsen bij hun eerste diagnostische hypothesen vrij grove categorieën hanteerden. Bovendien bleken zij, onafhankelijk van elkaar, gezamenlijk slechts een beperkt aantal categorieën nodig te hebben om alle hypothesen te kunnen indelen. De ingevulde kansen verschilden evenwel nogal (tabel). De hypothese 'carcinoom' (met name van het colon) werd door alle artsen genoemd, met kansen variërend van 10% tot 55%. Vooral de termen 'vermagering' en 'afvallen' hadden gefungeerd als triggermechanisme voor het opstellen van deze hypothesen. Andere hypothesen werden minder systematisch genoemd. De hypothesen 'inadequaate voedingspa-

troon' en 'psychosociale problematiek' werden door de meesten als één cluster gezien, omdat bij ouderen het eerste nogal eens een weerslag is van het tweede. Drie artsen vonden dat in dit cluster verreweg de meest aannemelijke oorzaak van het afvallen gelegen was. Twee artsen overwogen de diagnose 'enteritis', terwijl slechts één van de deelnemers de deur naar het alternatief 'geen afwijkingen' duidelijk openhield.

Opvallend is dat de meeste artsen zich wat de kansen betreft uitdrukten in veelvouden van 10%. Slechts een enkeling geeft een preciezere schatting. Opmerkelijk is ook dat de som van de kansen steeds 100% is. Blijkbaar acht men de kans dat twee of meer van de onderscheiden categorieën tegelijk aan de orde zijn, te verwaarlozen.

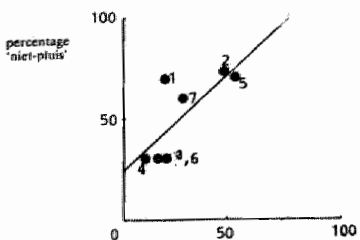
Er zijn dus aanmerkelijke verschillen in beoordeling wat de geschatte waarschijnlijkheid van de diverse diagnosecategorieën betreft. Het ligt voor de hand dat dit zich zal weerspiegelen in het verder te voeren beleid.

De volgende vraag komt nu naar voren: hoe groot is eigenlijk de kans dat iemand van 64 jaar coloncarcinoom heeft, of liever: hoe groot is deze kans bij een man van 64 jaar die sinds een jaar afvalt en passagère klachten van bloedverlies per anum en dunne ontlasting heeft gehad en nu het spreekuur van de huisarts bezoekt? Bij pogingen tot beantwoording blijkt dat naarmate men de schatting van de 'a priori-kans' voor meer specifieke situaties wil weten, het des te minder waarschijnlijk is dat deze kans op grond van verricht onderzoek te bepalen is of ooit precies te bepalen zal zijn. Subjec-

Tabel 1. Door de deelnemers geschatte kansen op diverse diagnose(categorieën) op grond van de gepresenteerde casus en hun 'niet-pluis'-gevoel.

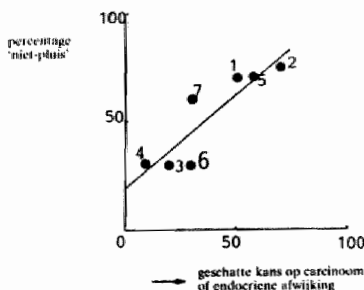
diagnose categorie	arts						
	1	2	3	4	5	6	7
carcinoom	20	50	20	10	55	20	30
inadequate voeding					40		
psychosociale problematiek	50	30	30		70	70	
endocrien: - diabetes	20						
- hyperthyroidie	20			3	10		
- hypothyroidie	10						
enteritis		50	30				
geen afwijking				60			
'niet-pluis'-gevoel	70	75	30	30	70	30	60

Figuur 1. Samenhang tussen 'niet-pluis'-gevoel en geschatte kans op carcinoom

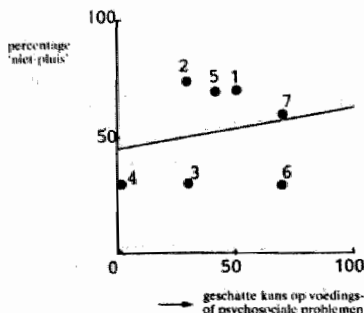


De punten 1 t/m 7 geven de schattingen van de artsen weer ( $r=0,75$ ).

Figuur 2. Samenhang tussen 'niet-pluis'-gevoel en geschatte kans op carcinoom of endocriene afwijkingen ( $r=0,89$ ).



Figuur 3. Samenhang tussen 'niet-pluis'-gevoel en geschatte kans op voedings- en/of psychosociale problemen ( $r=0,21$ ).



tieve schattingen, op grond van de eigen ervaring en die van anderen, zullen dan ook in de toekomst het handelen in belangrijke mate blijven sturen. Daarom zal uitwisseling van deze schattingen, en daardoor eventuele onderlinge aanvul-

ling of correctie ervan, niet kunnen worden gemist.

## 'PLUIS' EN 'NIET PLUIS'

Er bleken twee groepen artsen herkenbaar: een groep die de situatie overwegend als 'pluis' ervaarde en een groep die duidelijk neigde naar 'niet-pluis'. Een tussengroep (fifty-fifty) was er in dit geval niet (tabel).

Er is een duidelijk positieve samenhang tussen het 'niet-pluis'-gevoel en de geschatte kans op carcinoom (figuur 1) en, nog duidelijker tussen het 'niet-pluis'-gevoel en de geschatte kans op een carcinoom of een endocriene aandoening (figuur 2). Het 'niet-pluis'-gevoel hing praktisch niet samen met de geschatte kans op psychosociale problematiek en/of inadequate voeding (figuur 3).

Het onderscheid tussen 'pluis' en 'niet-pluis' wordt veel gehanteerd in het kader van de communicatie over de geschatte ernst van de situatie. Toch is het niet duidelijk wat hiermee precies wordt aangegeven en hoe de beoordeling tot stand komt. Opmerkelijk is nu dat veel artsen ook zonder duidelijkheid goed met dit onderscheid uit de voeten lijken te kunnen in hun communicatie met collega's. In de groep kwam wel de vraag naar voren of het onderscheid niet wat al te globaal was en aanleiding zou kunnen geven tot schijncommunicatie.

## STRUCTURERING PROBLEEM

Gevraagd werd of er deelnemers waren die een bepaald beleid wilden verdedigen. Het voordeel hiervan is dat men uitgesproken voorstanders van verschillende benaderingen met elkaar kan confronteren. Het ging vooral om de eerste stap die zou moeten worden genomen. Drie strategieën kwamen daarbij naar voren:

a. Verricht onmiddellijk colononderzoek. Daarbij kan dan worden gekozen tussen een röntgenonderzoek (X-colon) en een coloscopie.

b. Doe eerst een oriënterend laboratoriumonderzoek (met name BSE, Hb, bloedsuiker, faeces). Indien dit afwijkt, de waarden oplevert kan eventueel verdergaande diagnostiek plaatsvinden.

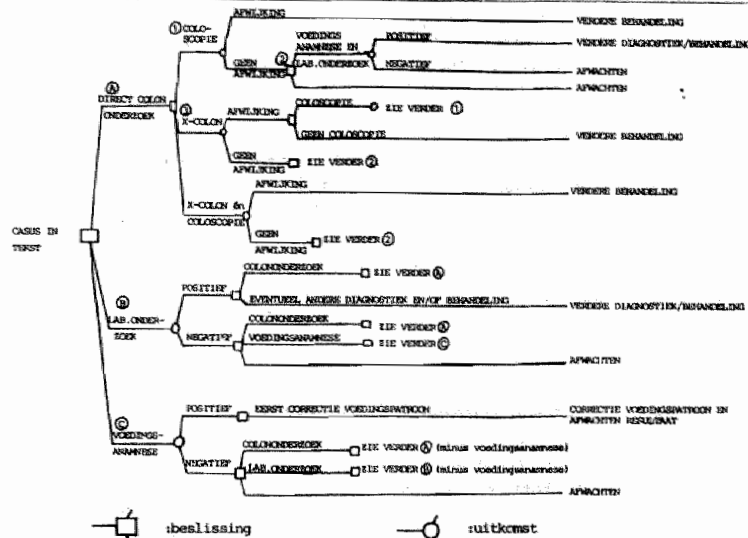
c. Begin met een nadere anamnese wat het voedingspatroon betreft. Levert deze nadere anamnese een verklaring op voor het afvallen, dan kan de aandacht in eerste instantie worden gericht op verbetering van het voedingspatroon.

Strategie C werd aangehangen door degenen die de kans op voedingsproblemen hoog inschatten (artsen 1 en 6). Strategie A werd gepropageerd door degenen die coloncarcinoom zo waarschijnlijk vonden dat colononderzoek toch zou moeten plaatsvinden, wat ook de uitslag zou zijn van ander diagnostisch onderzoek (artsen 2 en 5).

De weergave van strategieën is onvolledig als niet duidelijk is welke consequenties de uitkomsten van eerdere stappen kunnen hebben voor eventueel te nemen volgende stappen. Daarom is het vaak nodig eventueel volgende stappen te expliciteren. Het resultaat hiervan kan worden 'genotuleerd' in de vorm van het geraamte van een beslissboom waarin de belangrijkste opties zijn weergegeven. In dit geval ontstond *figuur 4*. Deze beslissboom bleek nog vrij ingewikkeld. De volgende *leerpunten* kwamen nu onder meer aan de orde in het kader van de discussie over de meest wenselijke strategie:

- Als een bepaalde diagnose erg waarschijnlijk is (hoge a priori-kans) zal de negatieve uitslag van een weinig sensitieve test niet gauw geruuststellen (bijvoorbeeld een normale BSE of een positieve voedingsanamnese bij een sterke verdenking op coloncarcinoom<sup>3</sup>).
- Een test waarvan de uitslag het beleid niet zal beïnvloeden, kan beter worden overgeslagen (zelfde voorbeelden<sup>4</sup>).
- Om uit te maken of men uit twee tests die in principe hetzelfde meten er één moet kiezen dan wel beide tegelijk moet doen of eventueel ná elkaar (en in welke volgorde) dient men een juiste indruk te hebben van sensitiviteit en specificiteit van beide tests. Ook moet men weten hoe sterk de uitslagen van beide tests met elkaar correleren c.q. van elkaar afhankelijk zijn (bijvoorbeeld X-colon en/of coloscopie<sup>5</sup>).
- Het is van belang overwegingen ten aanzien van stappen die pas verderop in de beslissboom aan de orde komen naar de primaire besluitvorming terug te koppelen. Bijvoorbeeld: als coloscopie een betere test zou zijn voor het aantonen of uitsluiten van coloncarcinoom dan een X-colon, dan pleit dit ervoor het X-colon over te slaan en direct coloscopie te verrichten. Het kan echter zijn dat men hiervoor terugschrikt, omdat men dit onderzoek in verhouding tot een X-colon zo belastend acht dat men het op een strengere indicatie wil laten verrichten; in voorkomende gevallen zal men dan toch eerst een X-colon aanvragen. In plaats

*Figuur 4. Voorlopige beslissboom naar aanleiding van de casus in de tekst, opgesteld tijdens een bijeenkomst met huisartsen in opleiding.*



van het aanvragen van een X-colon zou men wellicht ook eerst nog een andere optie willen proberen. Dit zou men zeker willen als beide colononderzoeken zo weinig valide zijn dat zij zowel veel carcinoomgevallen missen als veel niet-carcinoompatiënten ten onrechte de weg van nog verdergaand ingrijpen opsturen.

Indien er voldoende gegevens voorhanden zijn, zijn deze overwegingen te verwerken door middel van kwantitatieve analyse c.q. het 'doorrekenen' van de beslissboom. Indien men iedere mogelijke einduitkomst een bepaalde waarde (*utiliteit*) kan toekennen, kan men nagaan welke strategie de hoogste waarde bereikt<sup>6</sup>.

Dit dwingt ertoe de consequenties van (fout-)positieve en (fout-)negatieve uitslagen tegen elkaar af te wegen. Vaak is echter niet aan de voorwaarden voor een kwantitatieve analyse voldaan. Door een 'kwantitatieve' analyse van het besluitvormingsproces, waarbij dan vooral de structuur van de beslissboom en de identificatie van cruciale stappen van belang zijn, zal men dan vaak toch een zinvolle evaluatie van diverse strategieën kunnen verrichten. Ook kan men hieraan aangrijpingspunten voor nadere literatuurstudie of empirisch onderzoek ontleen (bijvoorbeeld over de relevante a priori-kansen en testeigenschappen). In vervolgbij-

eenkomsten kan het probleem dan verder worden uitgewerkt, ook met kwantitatieve gegevens.

#### SELECTIE CASUÏSTIEK

Een punt waaraan meer dan op het moment aandacht moet worden besteed is de selectie van te presenteren casuïstiek. Artsen kunnen geneigd zijn vooral succescasus in te brengen. Dat zijn bijvoorbeeld casus die hun voortreffelijk medisch kunnen demonstreren of een gepropageerde boodschap onderstrepen. Casus met een teleurstellende of beschamende afloop hebben, afhankelijk van de groep, een minimale of juist overdreven grote kans spontaan te worden uitgekozen. Casuïstiek waarvan men vindt dat zij weinig leerzaam bevat, bijvoorbeeld omdat zij verwarrend of alledaags is of een duidelijke afronding mist, wordt, zo leert de ervaring, zelden geselecteerd. Het resultaat is dat men bij de intercollegiale uitwisseling van ten opzichte van de dagelijkse praktijk sterk vertekend leer-materiaal uitgaat, waarin men zich niet meer herkent.

Er is nog een ander nadeel: omdat men achteraf graag de 'clou', de werkelijke afloop, vertelt of wil horen, krijgt deze de status van het juiste antwoord; dit kan een ten onrechte normerend effect met zich meebrengen. Ook is vertekening

vooraf mogelijk ten gevolge van selectie op onderwerp: in het kader van een volgens planning te bespreken onderwerp, bijvoorbeeld colonkanker, wordt aan één van de deelnemers gevraagd een eigen casus voor te bereiden. Het is duidelijk dat de bespreking in de groep in dit geval reeds van tevoren sterk is beïnvloed, zowel wat de te schatten a priori-kansen als de opties ten aanzien van het beleid betreft. Men ziet dit vaak gebeuren bij themagericht georganiseerde nascholing.

In het beschreven geval bleek in werkelijkheid een X-colon te zijn vervaardigd, met als uitslag: coloncarcinoom. Onbekend is evenwel hoe vaak bij vermagering een colononderzoek wordt verricht zonder dat kanker of een andere diagnose wordt vastgesteld. Het zou dus onjuist zijn naar aanleiding van deze casus zonder meer te concluderen dat bij een patiënt als in de casus altijd een colono foto of coloscopie moet worden gemaakt.

De zin van een casusbespreking is meer gelegen in de onderlinge uitwisseling naar aanleiding van de casus dan in de werkelijke afloop van de casus. Het be-

kend maken van de werkelijke afloop kan zelfs het leerproces achteraf ondermijnen en zou in principe achterwege kunnen blijven.

Wat de selectie van de te bespreken casuïstiek betreft is er veel voor te zeggen deze *at random* te ontleen aan de praktijkervaring van bijvoorbeeld de afgelopen week. Geen enkel geval zal daarbij oninteressant blijken, zelfs niet als het om algemeen ingeslepen routines gaat; integendeel: juist dan krijgt men de kans echte blinde vlekken op te sporen die verscholen liggen in een hele scala van dergelijke behandelingen.

#### SAMENVATTING

Aan de hand van het verloop van een bijeenkomst met huisartsen in opleiding werd beschreven hoe op eenvoudige wijze, reeds in één sessie, belangrijke aspecten van de medische besluitvorming aan de orde kunnen komen. Deze aspecten betreffen verschillen in beoordeling van kansen, medisch-besliskundige principes en het identificeren van cruciale stap-

pen in het beleid, alsmede van aangrijpingspunten voor nadere studie hierover. In de bespreking wordt de nadruk gelegd op het belang van een weloverwogen wijze van selectie van te presenteren casuïstiek.

★

★ ★

*Met dank aan de huisartsen in opleiding van groep 24, huisartsenberoepsopleiding Rijksuniversiteit Limburg.*

#### Literatuur

1. Neufeld VR, Barrows HS. The 'McMaster Philosophy': An approach to Medical Education. *Journal of Medical Education* 1974; 49: 1040-50.
2. Sackett DL. The critical appraisal of clinical evidence: strategies for its introduction, 1982. In: *Manual for a workshop on how to teach the critical appraisal of clinical evidence*. Department of clinical Epidemiology & Biostatistics, McMaster University, Hamilton, 1982.
3. Knottnerus JA. Interpretatie van de diagnostische gegevens. *Huisarts en Wetenschap* 1983; 26: 363-8.
4. Knottnerus JA, Leffers P. Het belang van therapeutische mogelijkheden voor diagnostische beslissingen. *The Practitioner* 1984; 1: 769-77.
5. Sturmans F. Evaluatie van het diagnostisch proces. *Syllabus PAOG*, 1984. Maastricht, 55-87.
6. Weinstein MC, Fineberg HV. *Clinical decision analysis*. Philadelphia: Saunders, 1980.

# Het belang van therapeutische mogelijkheden voor diagnostische beslissingen

J. A. KNOTTNERUS en P. LEFFERS

Capaciteitsgroep Epidemiologie/Gezondheidszorgonderzoek, Rijksuniversiteit Limburg, Maastricht

**D**E VUISTREGEL dat diagnostiek en behandeling niet los van elkaar gezien kunnen worden, is stokoud en tegelijk zeer actueel. Traditioneel bestaat de angst dat te weinig diagnostiek wordt verricht vóórdat therapeutische beslissingen genomen worden. De laatste jaren is echter ook steeds meer bezorgdheid ontstaan over het tegendeel. De diagnostiek is een eigen leven gaan leiden en is in bepaalde gevallen de therapeutische mogelijkheden boven het hoofd gegroeid.

Het gaat om twee kanten

**Samenvatting** *Het handelen van de huisarts berust in wezen op een voortdurende wisselwerking tussen zijn diagnostische activiteiten en de therapeutische overwegingen. Met een minimum aan feitenmateriaal, dat hem ter beschikking staat zonder de patiënt aan allerlei invasieve onderzoeken te moeten onderwerpen, zal de huisarts een aantal factoren tegen elkaar afwegen die de verdere gang van zaken sterk kunnen beïnvloeden.*

van dezelfde hulpverleningsmedaille: voor zinvolle behandeling is een optimale diagnostische onderbouwing nodig; en: diagnostiek heeft pas betekenis als haar uitkomsten kunnen bijdragen tot een betere hulpverlening.

## De carotissouffle

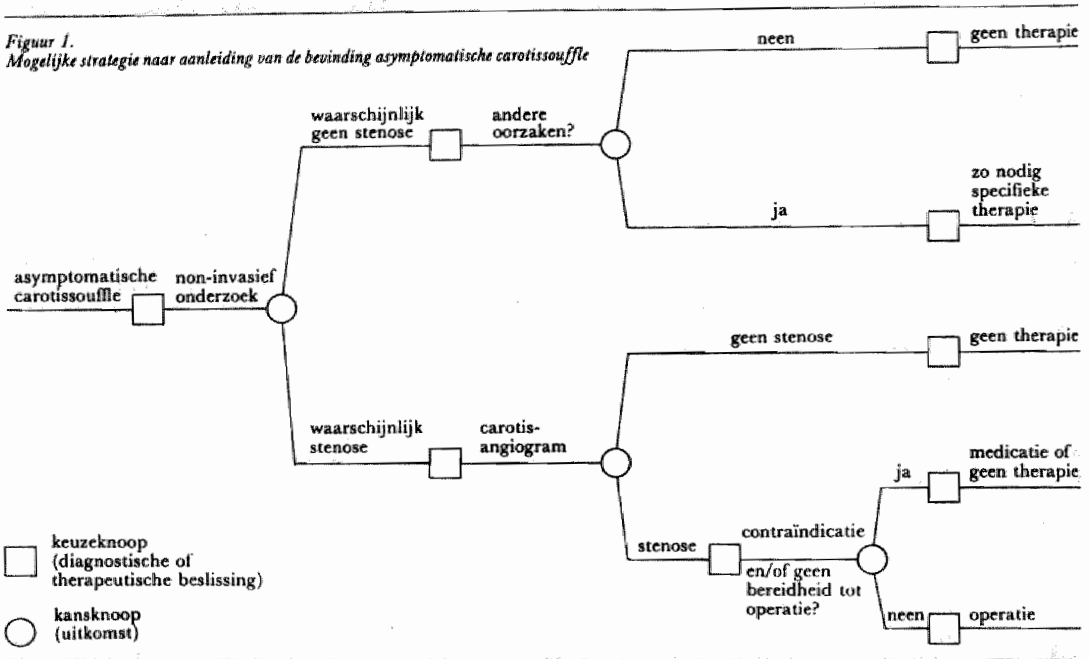
Als bij toeval, tijdens een keuring of bij een als routine uitgevoerd uitziftingsonderzoek, een soufflé in de hals wordt ontdekt, rijst de verdenking op een stenose van de A. carotis. Uitgaande van de opvatting dat zo'n stenose, ook bij klachtenvrije patiënten, een pathologisch proces inhoudt, ligt het voor de hand nader onderzoek te adviseren. Wellicht kan het euvel exact in kaart worden gebracht, de ernst ervan geschat, en misschien kan het verholpen worden met een operatie. Bovendien is voor de diagnostiek een groeiend aantal middelen beschikbaar: diverse toepassingen van Doppler en ultrageluid, oculoplethysmografie, digitale subtractie-angiografie en de beproefde carotis-

angiografie volgens Seldinger, om de belangrijkste te noemen. Men kan deze technieken bij iedere patiënt naast elkaar gebruiken om telkens een zo compleet mogelijk plaatje te krijgen. Van een meer kritische instelling getuigt het, als men te werk gaat volgens een bepaalde systematiek, een strategie, waarbij het beslissingsproces stapsgewijs wordt opgebouwd.

In figuur 1 is schematisch weergegeven hoe zo'n strategie voor het probleem van de asymptomatische carotissoufflé er uit kan zien.<sup>1</sup> Eerst wordt getracht door middel van niet-invasieve onderzoeksmethoden meer te weten te komen over een eventuele lokale oorzaak van de soufflé. Dit vergt in het algemeen een verwijzing. Vervolgens wordt, als er naar alle waarschijnlijkheid geen lokale stenose of obstructie is, nagegaan welke andere oorzaken er aan het symptoom ten grondslag kunnen liggen (zoals voortgeleide cardiale souffles of anemie) en aansluitend wordt beoordeeld welke speci-



Figuur 1.  
Mogelijke strategie naar aanleiding van de bevinding asymptotische carotissouffle



fieke therapeutische mogelijkheden daarvoor zijn. Is er op grond van het niet-invasieve onderzoek wel een stenose te verwachten, dan gaat men informatie verzamelen om te besluiten of er al dan niet geopereerd zou kunnen en moeten worden (endarteriëctomie). Met angiografie kan worden bepaald of er inderdaad een stenose bestaat, of deze operabel is, waar deze precies zit en hoe ernstig zij is. Voorts wordt

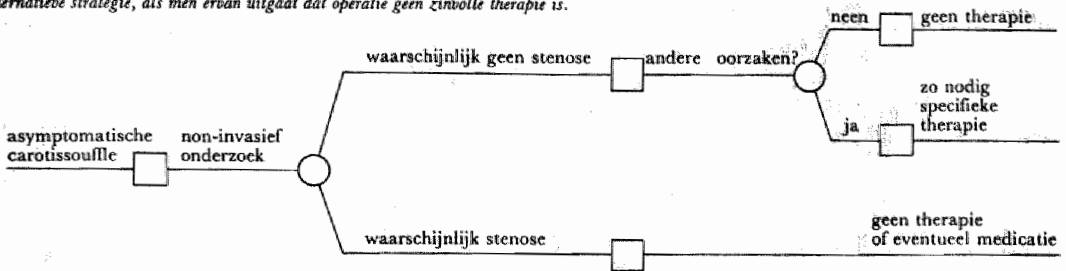
de vraag gesteld of er sprake is van contra-indicaties voor operatie, bij voorbeeld hoge leeftijd, een belangrijke longziekte of een ernstige inoperabele hartaandoening. Naar aanleiding van de beantwoording hiervan wordt ofwel besloten tot operatie, ofwel wordt de voorkeur gegeven aan een conservatieve benadering, al dan niet met behulp van trombocyten-aggregatieremmers.

Er is in elk geval sprake van

een complex probleem als men deze aanpak kiest. De gang van zaken kan vereenvoudigd worden, en bepaalde stappen kunnen worden kortgesloten.

– Als men zich al bij het vinden van de eerste symptomen rekenschap geeft van de eventuele contraïndicatie aan de orde zijn voor operatieve (of ook medicamenteuze) therapie, zal dikwijls aan invasieve diagnos-

**Figuur 2.**  
*Alternatieve strategie, als men ervan uitgaat dat operatie geen zinvolle therapie is.*



tiëk – die alleen nodig is om de operatie-indicatie definitief te stellen – niet meer gedacht hoeven te worden. Men kan zich dan bij veel mensen, bij wie anders wellicht angiografie verricht was, direct beperken tot de behandeling van, vaak al bekende, andere gezondheidsproblemen. Eventueel kan een niet-invasief onderzoek zinvol zijn om de indicatie tot het instellen van medicamenteuze therapie voldoende te documenteren. In andere gevallen, bij voorbeeld bij een ernstig cardiovasculair lijden waarbij reeds stollingwerende middelen of trombocyten-aggregatierepressoren gegeven worden, had men bij wijze van spreken al op voorhand kunnen afzien van auscultatie van de carotiden.

– Bij het bespreken met de pa-

tiënt van de betekenis van de bij fysisch onderzoek gevonden soufflé, kan al direct duidelijkheid worden verkregen over de vraag of deze überhaupt tot een eventuele operatie bereid zou zijn. Dit is van groot belang, want als die bereidheid er niet is, kan veel diagnostiek achterwege blijven.

– Een nog sterkere vereenvoudiging kan optreden, als men weet dat het nut van operatie bij asymptomatische<sup>2,3</sup> en zelfs bij de symptomatische<sup>4</sup> carotisstenosen nooit afdoende is aangetoond (fig. 2). Als er al enig nut zou zijn, is bijzonder twijfelachtig of dit wel opweegt tegen de nadelen en risico's van invasieve diagnostiek en operatie. En ook de zin van medicinale behandeling staat niet buiten kijf.<sup>5</sup> Als men op dit spoor

doordent, zou auscultatie van de carotiden alleen zinvol zijn om de patiënt beter te kunnen informeren over zijn of haar toestand. Zelfs dan is nog de vraag of men de patiënt goed doet met dergelijke mededelingen.

### **Wisselwerking diagnostiek-behandelingsplan**

De boodschap is duidelijk: het denken over de (mogelijkheden van) behandeling en over de consequenties daarvan voor de mogelijk te vergaren informatie, dient vanaf de aanvang van de hulpverlening de aandacht te hebben. Het gehele beleid kan dan stap voor stap worden ontwikkeld. Dreigende overmaat kan worden herkend en voorkomen.

Daarbij is sprake van een voortdurende wisselwerking tussen de diagnostische activiteit en therapeutische overwe-

gingen. Een al te gefascerde aanpak heeft gevaren. De houding 'eerst maar eens rustig de diagnostiek afwerken, de therapie komt daarna aan de beurt' kan, bij voorbeeld, leiden tot myelografieën bij mensen die toch niet aan hun hernia geopereerd hadden willen worden, of tot uitvoerige tumordiagnostiek bij patiënten die al ten dode zijn opgeschreven.

Van iedere beslissing moet dus tevoren het nut voor de patiënt worden beoordeeld. Wat betekenen de mogelijke uitkomsten voor mijn verdere beleid? Brengt de mogelijk uitkomst mij verder dan ik nu ben? Daarbij dienen ook de voorkeuren van de patiënt, diens conditie, en vooral ook de merites van de behandelingsalternatieven in de overweging te worden betrokken.

Kortom: de klachten of symptomen waarmee de patiënt zich aandient, leiden tot de vorming van diagnostische hypothesen.<sup>6</sup> Of en hoe deze getoetst dienen te worden hangt mede af van de therapeutische overwegingen die bij de hypothesen relevant zijn.

In de vroegste fase van de hulpverlening, als er nog weinig of geen medische activiteit in gang is gezet, vinden we de huisartsgeneeskundige betrok-

kenheid. Juist de huisarts zal dan ook afwegingen moeten maken, die bepalend zijn voor de verdere gang van zaken. Dat betekent dat de huisarts, over een breed terrein, erg goed moet weten wat de geneeskunde wel of niet te bieden heeft. Dat vergt een specifieke deskundigheid die tijdens de opleiding maar mondjesmaat wordt aangeboden, en vaak later verworven moet worden. Hier ligt ook een belangrijk aangrijpingspunt voor de ontwikkeling van het overleg tussen huisartsen en specialisten over geneeskundige thema's of, via consultatie, over concrete gevallen.

Onder dit hoofdje willen we nog wijzen op de gevaren van diagnostische routinepakketten: 'Bij patiënten met die-en-die klachten beginnen we altijd met die-en-die onderzoeken'. Enerzijds kunnen dublures voorkomen, en overmatige belasting van de patiënt. Anderzijds brengt het inwinnen van meer informatie vaak ook een grotere kans op foute of moeilijk te interpreteren gegevens met zich mee. Bovendien verleert men door al te routinematig handelen het analytische denken met kritische evaluatie van iedere te nemen of genomen stap.

### **Kunnen weten is niet: moeten weten**

Het is niet moeilijk om elk gevoel van onzekerheid om te zetten in een streven van maximaal haalbare zekerheid. Vaak leidt dit echter tot hulpverlening van twijfelachtige kwaliteit. Een veel grotere en meer wezenlijke kundigheid vergt het om, nu eenmaal bestaande, onzekerheden als uitgangspunt te aanvaarden, en telkens te bezien in hoeverre het zinvol is om tot meer zekerheid te geraken. Zoals Reuben het uitdrukt: 'Ability to know' moet niet verward worden met 'need to know'.<sup>7</sup> Vaak is een bescheiden aantal, soms maar één of twee, bevindingen al zo cruciaal dat additionele informatie er weinig meer toe doet. Zo is bij de klacht pijn op de borst de lokalisatie, het karakter van de pijn en of deze bij inspanning optreedt, al zo essentieel dat bij een op deze punten niet-verdachte anamnese uitvoerig nader onderzoek volgens velen beter achterwege kan blijven.<sup>8</sup> En bij acute pijn in de rechter onderbuik valt het belang van bloedonderzoek zodanig in het niet bij dat van de fysische diagnostiek, dat het bestaansrecht ervan in de primaire besluitvorming op zijn minst twijfelachtig is.

Voor al de huisarts is ge-

wend om met de methode van de 'cruciale' bevindingen te werken, en dit is ook een noodzaak gezien zijn positie in de gezondheidszorg en het klachtenaanbod dat hij ontmoet. Wellicht verdient de verdere ontwikkeling van deze 'methode' in de nabije toekomst minstens zoveel aandacht als het wetenschappelijk onderzoek op het gebied van de geavanceerde diagnostiek.

Een belangrijke voorwaarde om tot een meer terughoudende geneeskunde te geraken, is dat een dergelijke geneeskunde als goede vorm van beroepsuitoefening aanvaard wordt, zowel binnen de medische wereld als door de patiënten. In dit verband is het, bij voorbeeld, tekenend dat op dit moment in de rubricering van de beslissingen van de medische tuchtcolleges door de Geneeskundige Hoofddirectie wél voorkomen: niet of te laat komen op verzoek; onvoldoende zorg; onvoldoende informatie; en niet of te laat verwijzen. En niet de rubrieken: ongevraagd of te veel komen; veel te veel zorg; een overmaat aan informatie; ten onrechte, te vroeg, of te veel verwijzen; en ten onrechte niet, te laat of te weinig terugwijzen naar de huisarts. En ook het begrip 'kunstfout' is nog eenzijdig gekleurd.

### Principes

Een probleem bij het gebruik van bepaalde diagnostica is, of de te verkrijgen extra informatie opweegt tegen de belasting van de patiënt. Maar ook het steeds minder belastend worden van allerlei diagnostische activiteiten, bij voorbeeld door het vervangen van invasieve door niet-invasieve technieken, kan, hoe prettig dit op zichzelf ook is, een eigen problematiek met zich meebrengen. Men krijgt de neiging ze eerder te gebruiken, ook al is de klinische relevantie van dat gebruik niet altijd even duidelijk.

Het ontdekt worden van een diagnosticum is soms al voldoende reden om het te gebruiken, en als er één in diskrediet raakt, staan er al nieuwe klaar waarvan gezegd wordt dat de bezwaren niet meer gelden. Er is veel voor te zeggen om voor introductie van diagnostica in de praktijkvoering een procedure in te voeren die vergelijkbaar is met die welke geldt voor geneesmiddelen. Algemene acceptatie zou pas moeten plaatsvinden als door middel van goed onderzoek aangetoond is dat datgene wat gepretendeerd wordt ook klopt. Tevens zou de vraag moeten worden beantwoord wat de nieuwe techniek toevoegt aan de bestaande mogelijkheden.

Wij besluiten met het weergeven van enkele basale principes die bij diagnostisch onderzoek van belang zijn, en sluiten ons daarin aan bij Reuben:<sup>7</sup>

- Als de uitslag van een onderzoek een verandering kan opleveren van het therapeutische beleid, dan moet dat onderzoek gedaan worden dat de noodzakelijke informatie oplevert met een minimum aan belasting en risico's voor de patiënt.
- Een onderzoek kan weliswaar niet van belang zijn voor het bepalen van de therapeutische strategie, doch geïndiceerd zijn om andere redenen: inzicht bieden in de prognose, of ter geruststelling van de patiënt. Ook dan moet men echter positieve argumenten hebben voor het onderzoek.
- Men moet onderscheid kunnen maken tussen diagnostiek ten behoeve van de individuele patiënt en diagnostiek ten behoeve van wetenschappelijk onderzoek. Beide hoeven niet samen te vallen. In het tweede geval is op zijn minst een goed research-protocol nodig, alsmede 'informed consent' van de patiënt. □

## Literatuur

1. Hurst, J. W.; Hopkins, L. C.; Smith, R. B.: *Noises in the neck*. The N. Engl. J. of Med., 1980; 302: 862-863.
2. Strandness, D. E.: *Why the controversy about carotid endarterectomy?* J. Cardiovascular Med., 1981; 6: 1067-1068 (Editorial).
3. Fields, W. S.: *The asymptomatic carotid bruit - operate or not?* Current Concepts of Cerebrovascular disease. Stroke, 1978; XIII: 1-14.
4. Warlow, C.: *Transient Ischemic Attacks*. In: Matthews, W. B.; Glaser, G. H.: *Recent advances in clinical neurology*. Edinburgh, Churchill Livingstone, 1982: 191-214.
5. Sorensen, P. S.; Pedersen, H.; Marquardsen, J. et. al.: *Acetylsalicylic acid in the prevention of stroke in patients with reversible cerebral ischemic attacks*. Stroke, 1983; 14: 15-22. A Danish cooperative study.
6. Knottnerus, J. A.; Sturmans, F.: *Het vormen van toetsen van diagnostische hypothesen*. The Practitioner, 1984; 1: 277-281.
7. Reuben, D. B.: *Learning diagnostic restraint*. The N. Engl. J. of Med., 1984; 310: 591-593.
8. Todd, J. W.: *Query cardiac pain*. The Lancet, 1983; august 6: 330-332.

---

DE ASYMPTOMATISCHE CAROTISSTENOSE, WEL OF NIET OPEREREN?

In het kader van een algemeen onderzoek ontdekt de huisarts door middel van auscultatie bij een patiënt een soufflé over een halsslagader welke mogelijk wordt veroorzaakt door een carotisstenose. Hij wordt nu met enkele belangrijke vragen geconfronteerd. De belangrijkste zijn:

- 1) heeft de patiënt nu een verhoogd risico op een cerebrovasculair accident? en:
- 2) is therapeutisch ingrijpen zinvol?

De eerste vraag kan positief worden beantwoord. Mensen met een carotisstenose hebben een 2 à 3 keer zo grote kans op cerebrovasculaire complicaties in de toekomst. Daarmee is vraag 2 nog niet beantwoord. Juist omdat patiënten met een carotisstenose vaak geen klachten hebben, is de vraag naar de consequenties van deze bevinding extra belangrijk. Wat heeft de literatuur ons hierover te melden? Het nu volgende artikel gaat hierop in.

DE ASYMPTOMATISCHE CAROTISSTENOSE, WEL OF NIET OPEREREN?

- een besliskundige benadering

A. Knottnerus <sup>1)</sup> - P. Leffers <sup>1)</sup> - J. Lodder <sup>2)</sup> - F. Sturmans <sup>1)</sup>

Samenvatting.

Met behulp van besliskundige technieken en gepubliceerde gegevens zijn de te verwachten klinische uitkomsten geëvalueerd van operatieve en conservatieve behandeling van asymptomatische carotisstenose, na respectievelijk 1, 4 en 8 jaar follow-up. Bij conservatieve behandeling is de mortaliteit het laagst, en de kans op TIA het grootst, onafhankelijk van de follow-up duur. De kans op een CVA is na één jaar het laagst bij conservatieve behandeling, doch na 4 en 8 jaar bij operatie. De verwachte utiliteit is steeds hoger voor conservatieve behandeling, en dit resultaat is stabiel over een breed traject van afzonderlijke kansen en utiliteiten.

Het opereren van asymptomatische carotisstenosen kan op grond van de literatuur niet gerechtvaardigd worden.

trefwoorden:   asymptomatisch/carotis/stenose/behandeling/besliskunde.

1) capaciteitsgroep Epidemiologie/Gezondheidszorgonderzoek

2) capaciteitsgroep Neurologie

## Inleiding.

Over het te volgen beleid naar aanleiding van de bevinding "asymptomatische carotissouffle" wordt verschillend gedacht (1). De kernvragen zijn: 1) of endarterectomie in geval van een eventuele carotisstenose de prognose verbetert en dus ook of 2) diagnostiek, gericht op het stellen van een operatie-indicatie, verricht moet worden.

Sommigen menen dat vanwege de verhoogde kans op een ischemisch accident een op operatie gericht beleid gewenst is (2,3). Volgens anderen slaat de balans van kosten (directe en indirecte risico's van angiografie en operatie) en baten (verbetering van de prognose door het opheffen van de stenose) door naar de conservatieve benadering (1,4,5). In de twee studies, waarin de klinische resultaten van geopereerden en niet-geopereerden met elkaar vergeleken werden, kwamen de geopereerden er beter af (6,7). Echter, de toewijzing van de behandeling gebeurde niet at random, maar werd bepaald door indicaties en contra-indicaties. De prognose van de geopereerden is daardoor waarschijnlijk a priori al gunstiger geweest. Een ander probleem is dat onderzoeksresultaten in zulke uiteenlopende vormen gepresenteerd worden, dat onderlinge vergelijking vaak moeilijk is. Resultaten van studies met een follow-up van 2 tot 15 jaar, en met een follow-up van minstens 6 maanden zijn niet zo maar met elkaar te vergelijken. Uiteindelijk zal de controverse alleen op te lossen zijn door middel van een gecontroleerd experiment op basis van randomisatie (8,9,10).

Zo'n experiment is tot nu alleen gedaan bij symptomatische patiënten (11).

Inmiddels moeten in de praktijk beslissingen worden genomen. Het ligt voor de hand dat men zich daarbij mede baseert op gepubliceerde gegevens. In het volgende hebben wij beslistkundige technieken gebruikt om beschikbare literatuurgegevens op systematische wijze te verwerken.

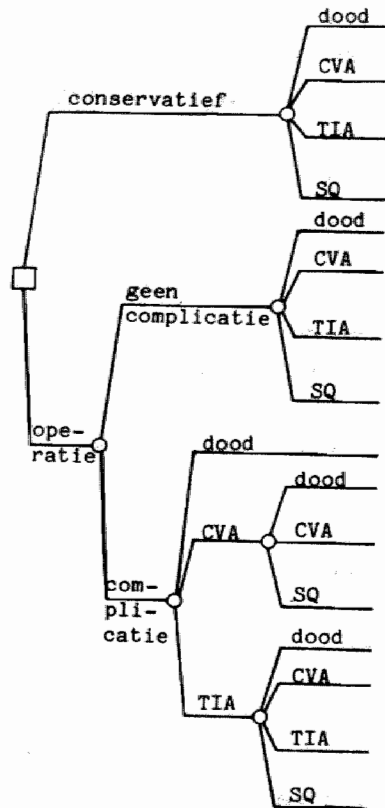
## Methode.

Structurering van het probleem.

Diagnostiek gericht op het vaststellen van operabele stenosen, heeft alleen zin als operatie van zo'n stenose effectief is. De kern van het



probleem is daarom de vergelijking van de resultaten van endarterectomie en conservatieve behandeling. Deze is ondergebracht in een beslisboom (figuur 1). Als uitkomsten zijn gekozen: sterfte, cerebrovasculair accident met blijvende gevolgen (CVA), transient ischemic attack (TIA), en de situatie waarin geen van deze gebeurtenissen optreedt (status quo).



Figuur 1: Beslisboom voor de behandeling van asymptomatische carotisstenose.

SQ = Status Quo  
□ : Keuzeknoop  
○ : Kansknoop

Kansen.

Voor het evalueren van een beslisboom dient men de kansen op de diverse klinische uitkomsten te kunnen berekenen (12). De kansen die in de beslisboom werden gebracht, zijn ontleend aan de literatuur over angiografisch aangetoonde asymptomatische stenosen. Hierbij moet opgemerkt worden dat van de 20 gevonden publicaties over het beloop van asymptomatische stenosen bij conservatief en chirurgische beleid (en perioperatieve complicaties) er 11 afvielen vanwege ontbreken van relevante gegevens of ernstige kwalitatieve tekortkomingen. Wat betreft het beloop na ongecompliceerde operatie en bij conservatieve behandeling gingen wij er vanuit dat de klinische verwickelingen gelijkmatig gespreid zijn over de follow-up periode (13-17). De betreffende kansen uit elke daarvoor geschikte studie konden nu worden berekend per tijdseenheid, door het aantal opgetreden klinische verwickelingen te delen door de som van de follow-up tijd van alle patiënten. Bij gebrek aan gepubliceerde gegevens over het klinisch beloop na het optreden van niet-fatale operatiecomplicaties, gebruikten wij hiervoor dezelfde kansen als na ongecompliceerde operatie. Dit kan de analyse hebben beïnvloed ten voordele van operatie.

Het eventuele effect van trombocyten-aggregatie remmers werd buiten beschouwing gelaten omdat hierover geen studies bij asymptomatische patiënten zijn gepubliceerd, en omdat in de meeste patiëntenseries informatie over eventuele medicatie ontbrak. Met geslachtsverschillen hebben wij geen rekening kunnen houden omdat specifieke cijfers hierover in de literatuur ontbraken. Gezien de gemiddelde leeftijd in de gepubliceerde series mag de analyse van toepassing worden geacht op oudere patiënten. Omdat de niet-geopereerde patiënten gemiddeld wat jonger waren dan de geopereerden (gemiddeld respectievelijk 63 en 66 jaar oud) corrigeerden wij het sterftecijfer van de eerste groep met behulp van de leeftijdsspecifieke sterftecijfers van de algemene bevolking (zie appendix).

Bij het doorrekenen van de beslisboom werd uitgegaan van rekenkundige gemiddelden van de aan de diverse publicaties ontleende kansen. Iedere aparte studie kan immers gezien worden als een steekproef uit de klinische ervaring ter plaatse, en er is dan geen reden om grotere studies een zwaarder gewicht toe te kennen.

## Utiliteiten.

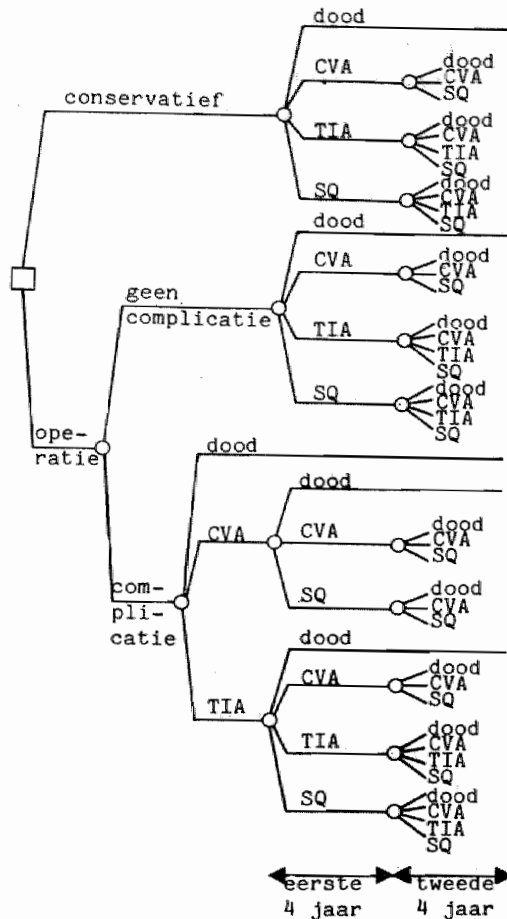
Voor een besliskundige analyse is het van belang dat aan iedere mogelijke klinische uitkomst een waarde wordt toegekend, een zogenaamde utiliteit. Deze waarde dient de vermindering van zowel de levensverwachting als de kwaliteit van leven te weerspiegelen. De verwachte waarde per strategie wordt dan berekend door de kans op elke uitkomst te vermenigvuldigen met haar utiliteit, en deze producten voor de betreffende strategie bij elkaar op te tellen (18).

Als utiliteiten voor sterfte, CVA, TIA en "status quo" kozen wij in eerste instantie respectievelijk 0.0, 0.75, 0.95 en 1.0. Deze factoren geven aan hoe groot de waarde van het verkeren in de betreffende toestand geacht wordt ten opzichte van gezond zijn. Daarbij gaat het om gemiddelden, waarin bijvoorbeeld voor het CVA zowel zeer ernstige als zeer lichte gevallen zijn verdisconteerd. Omdat deze keuzen arbitrair zijn, pasten wij hierop een sensitiviteitsanalyse toe. Hierbij gingen wij na, in hoeverre het resultaat van de berekening "gevoelig" was voor verandering van de utiliteiten (12).

## Lange termijn resultaten.

Omdat de gemiddelde follow-up duur van de gebruikte onderzoeken 4 jaar bedroeg, berekenden wij behalve de resultaten na 1 jaar ook die na 4 jaar follow-up. Voorts werd de toestand na 8 jaar geanalyseerd met behulp van een uitgebreide beslisboom (figuur 2).

Omdat cerebrovasculaire complicaties zich kunnen herhalen werden hierbij ook gegevens betrokken over de kans op sterfte, CVA en TIA bij patiënten die reeds eerder een CVA en/of TIA hadden gekregen. Voor patiënten die in de eerste periode symptomvrij bleven werden in de tweede periode van 4 jaar dezelfde kansen gebruikt als in de eerste periode (15). Voor het leeftijdsverschil van 4 jaren ten opzichte van de eerste periode werden de sterftekansen gecorrigeerd (zie appendix). De resultaten van de analyse werden uitgedrukt in "Quality Adjusted Life Years" (QUALY's): Deze uitkomstmaat is een aanduiding voor het aantal geleefde jaren, gecorrigeerd voor de "kwaliteit" van de gezondheidstoestand gedurende die jaren (12, 18).



Figuur 2: Beslisboom voor de behandeling van asymptomatische carotisstenose, voor de analyse van de situatie na 8 jaar follow-up.

SQ : Status Quo : geen verandering van de situatie.

Voor iedere mogelijke "patiëntencarrière" (overeenkomend met iedere mogelijke route die in de beslisboom is aangegeven) werd het aantal QUALY's berekend door het aantal jaren dat men in een bepaalde toestand verkeerde te vermenigvuldigen met de utiliteit van die toestand, en de aldus verkregen producten voor die patiëntencarrière bij elkaar op te tellen. Voor de utiliteit van een nieuwe toestand na een reeds eerder opgetreden verwikkeling is het product van de betrokken utili-

teiten genomen. Daarbij werd er vanuit gegaan dat perioperatieve complicaties na 0 jaar optraden, en de latere verwickelingen gemiddeld halverwege de eerste respectievelijk tweede periode.

Dus voor een patiënt die post-operatief een TIA en vervolgens in de eerste 4 jaar een CVA kreeg, om tenslotte in de tweede 4 jaars-periode te overlijden wordt in 8 jaren follow-up het volgende aantal QUALY's geteld:

in de eerste periode van 4 jaar:  $(2 \times 0.95) + (2 \times 0.95 \times 0.75)$

in de tweede periode van 4 jaar:  $\frac{(2 \times 0.95 \times 0.75)}{4.75 \text{ QUALY's}} +$   
totaal:

Omdat de lezer zijn eigen oordeel kan hebben over utiliteiten hebben wij ook aangegeven hoe groot de kans is, dat men na een bepaalde periode in een bepaalde toestand verkeert.

#### Sensitiviteitsanalyse.

Omdat het resultaat van de analyse afhangt van de gehanteerde kansen en utiliteiten, werd de waarde van elke strategie herberekend bij andere keuzen voor de kansen en utiliteiten. Hierdoor werd een indruk verkregen van de mate waarin het resultaat gevoelig is voor variatie van kansen en/of utiliteiten (12).

De berekeningen werden uitgevoerd op een Apple II<sup>e</sup> microcomputer, met het programma Decision Maker (19).

#### Resultaten.

De resultaten van het literatuuronderzoek zijn gegeven in de tabellen 1 en 2. De uit de verschillende publicaties afgeleide kansen liggen nogal uiteen, hetgeen kan worden toegeschreven aan verschillen wat betreft selectie van patiënten, kwaliteit van zorg en aan het toeval. Ten aanzien van het beloop na ongecompliceerde operatie stemmen de diverse studies nog het meest overeen. De resultaten van conservatieve behandeling zijn waarschijnlijk in ongunstige zin beïnvloed omdat in de meeste onderzoeken uitgegaan is van patiënten die in het verleden een contralaterale endarterectomie hadden moeten ondergaan (tabel 1).

Tabel 1: Mortaliteit en incidentie van CVA en TIA per periode van 4 jaar voor conservatief beleid, na geslaagde operatie, respectievelijk na het optreden van CVA en TIA. Literatuurgegevens.

	refe- renties	gemidd. leef- tijd	gemidd- deld	mortaliteit range	gecorr.* gemidd.	CVA gemidd- deld	CVA range	TIA gemidd- deld	TIA range	opmerkingen
BELOOP BIJ CONSERVA- TIEVE BE- HANDELING	6,15, 20-22	63	0,256	0,180-0,384	0,279*	0,064	0,016-0,160	0,224	0,152-0,284	In 4 studies was in het verleden contralateraal geopereerd.
BELOOP NA GESLAAGDE OPERATIE	6,23, 24	66	0,356	0,284-0,404	0,356	0,024	0,020-0,052	0,040	0,040	Slechts in één studie incidentie van TIA gemeld
BELOOP NA OPTREDEN CVA	26-28	67	0,376	0,320-0,464	0,408**	0,304	0,304	-	-	Slechts in één studie incidentie recidief-CVA gemeld.
BELOOP NA OPTREDEN TIA	9, 29-33	60	0,221	0,120-0,320	0,302**	0,167	0,040-0,320	0,459	0,120-0,712	

Gecorrigeerd is voor verschillen in gemiddelde leeftijd ten opzichte van de geopereerde patiëntengroep\* en voor het leeftijdsverschil van 4 jaar tussen de eerste en de tweede periode\*\*, met gebruikmaking van de leeftijdspe-  
cifieke sterftecijfers van de algemene bevolking (zie appendix).

Tabel 2: Peri-operatieve mortaliteit en morbiditeit van endarterectomie wegens asymptomatische carotis-stenose, in percentages van het aantal ingrepen.

referenties	complicatie	gemiddeld	range
6,17,23-25	mortaliteit	0,3%	0 - 1,5%
6,17,23-25	CVA	4,5%	0 - 18,2%
6,17,24	TIA	3,6%	1,2 - 7,1%

Tabel 3 bevat de resultaten van de analyse na respectievelijk 1, 4 en 8 jaar follow-up. De kans op sterfte is na elke follow-up duur het laagst bij conservatieve behandeling. De kans op een niet-fataal CVA is na 1 jaar het hoogst bij operatieve behandeling, doch na 4 en 8 jaar bij conservatief beleid. De kans op TIA is steeds het laagst bij operatief ingrijpen. De verwachte utiliteit is na elke follow-up periode het hoogst voor conservatieve therapie.

Tabel 3: Berekenende kansen op diverse uitkomsten na 1, 4 en 8 jaar follow-up voor de operatieve (OP) en conservatieve (CONS) strategie, en de te verwachten utiliteiten van beide strategieën.

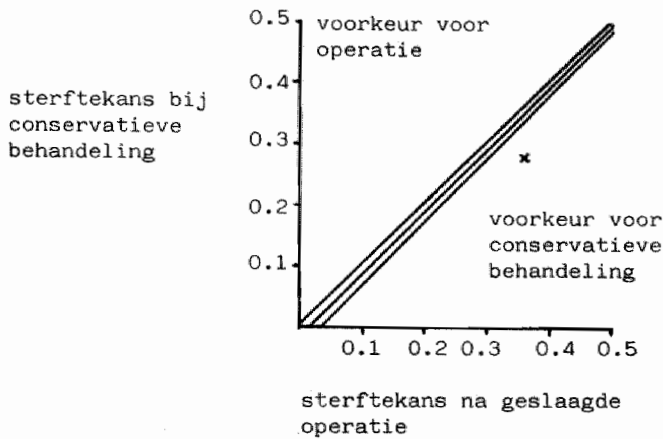
uitkomst	kansen na 1 jaar		kansen na 4 jaren		kansen na 8 jaren	
	OP	CONS	OP	CONS	OP	CONS
sterfte	0,092	0,070	0,358	0,279	0,609	0,511
CVA	0,047	0,016	0,052	0,064	0,051	0,103
TIA	0,041	0,056	0,059	0,224	0,054	0,216
geen neurologische symptomen	0,820	0,858	0,531	0,433	0,287	0,170
verwachte utili- teit *	0,939	0,963	3,22	3,39	5,22	5,66

\* De utiliteiten zijn uitgedrukt in voor kwaliteit gecorrigeerde levensjaren (QUALY's). Een jaar CVA- respectievelijk TIA-patiënt zijn is equivalent aan 0.75 respectievelijk 0.95 jaar vrij zijn van neurologische symptomen.

Blijkens de sensitiviteitsanalyse is dit resultaat niet gevoelig voor extreme variatie van de utiliteiten. Of men nu de "utiliteit" van een CVA op 0.0 of 1.0 stelt, de conservatieve behandeling blijft de voorkeur houden (tabel 4).

Tabel 4. Verwachte utiliteit van operatieve (OP) en conservatieve (CONS) strategie na 1, 4 en 8 jaren follow-up bij extreme variatie van de utiliteit van het CVA.

follow-up duur	utiliteit van CVA = 0.00 TIA = 0.95		utiliteit van CVA = 1.0 TIA = 1.0	
	OP	CONS	OP	CONS
1 jaar	0,920	0,954	0,947	0,965
4 jaren	3,07	3,29	3,32	3,44
8 jaren	4,92	5,32	5,34	5,86



Figuur 3: Sensitiviteitsanalyse met variatie van de sterftecijfers bij conservatieve behandeling en na geslaagde operatie van asymptomatische carotisstenose. Toestand na 8 jaar follow-up. De sterftecijfers zijn uitgedrukt als sterftekans per periode van 4 jaar follow-up. De voorkeur voor operatie betreft alle punten (het gebied) boven de lijnen, en de voorkeur voor conservatieve behandeling komt overeen met het gebied eronder. De lijnen komen overeen met kansen op operatiecomplicaties van respectievelijk 15% (bovenste lijn), 8.4% (middelste lijn) en 0% (onderste lijn).

x : puntschatting op grond van de literatuur.

Ook wat betreft de gebruikte kansen is het resultaat vrij stabiel. Uit figuur 3 blijkt in hoeverre onze schattingen van de mortaliteit bij conservatieve behandeling, en na ongecompliceerde operatie, onjuist dienen te zijn opdat alsnog de operatieve strategie de voorkeur zou verdienen. Hiertoe zou de mortaliteit bij conservatieve benadering



aanzienlijk ongunstiger moeten worden ten opzichte van de mortaliteit na ongecompliceerde operatie, en wel des te meer naarmate de kans op perioperatieve complicaties toeneemt. De figuur betreft de situatie na 8 jaar, maar na 1 en 4 jaar is vrijwel hetzelfde beeld te zien. Variatie van andere kansen heeft nog veel minder invloed dan de twee genoemde sterftcijfers.

### Discussie

Op grond van de gepubliceerde gegevens kan blijkens onze analyse een op operatie gericht beleid niet worden verdedigd. Na operatie lijkt de sterfte over de follow-up periode hoger te zijn. Dit wordt, ook bij extreme variatie van de utiliteiten van CVA en TIA, niet gecompenseerd door een lager risico op CVA en TIA. Deze conclusie kan getrokken worden zonder dat nog rekening gehouden wordt met de eventuele effecten van plaatjesremmers.

De mogelijke vertekeningen in de analyse werken vooral ten gunste van operatie, zodat in werkelijkheid het voordeel van de conservatieve strategie wellicht nog groter is. Immers, het ligt voor de hand dat de prognose van geopereerden a priori beter was omdat bij hen geen contra-indicaties voor operatie waren vastgesteld. Bovendien doet het feit, dat in de conservatief behandelde groepen in het verleden vaak wel contralateraal geopereerd was, vermoeden dat het hier ging om patiënten met een vrij uitgebreid lijden. Voorts zijn in de analyse de risico's van angiografie buiten beschouwing gelaten. Hoewel aangenomen mag worden dat deze risico's tegenwoordig gering zijn (3, 34, 35), zullen ze toch vooral op rekening komen van een beleid waarbij operatie wordt overwogen. De belasting ten gevolge van angiografie en operatie op zichzelf, dus los van de eventuele complicaties, zou men nog als extra nadeel aan de operatieve strategie kunnen toevoegen. Ook moet men ermee rekening houden, dat onderzoekers liever gunstige dan ongunstige behandelingsresultaten publiceren. Dit laatste kan echter ten aanzien van zowel geopereerden als niet-geopereerden een rol spelen.

Verondersteld mag worden dat de risico's van operatie lager zijn in gespecialiseerde chirurgische centra met veel ervaring. Echter, ook

als de kans op operatie-complicaties op 0 gesteld wordt, blijven de resultaten na operatie ongunstiger. In de literatuur varieert het totale percentage operatiecomplicaties (mortaliteit en morbiditeit tesamen) dat men nog acceptabel acht van 1 tot 3% (6, 8, 36). De laatste 10 jaar zijn de operatie-risico's afgenomen als gevolg van de introductie van nieuwe technieken (37, 38). Het is nog de vraag of hierdoor ook de lange termijnprognose zal verbeteren. Bovendien worden er nog steeds complicatiepercentages gerapporteerd die hoger liggen dan die welke wij in onze analyse (tabel 3) gebruikten (39).

In de afweging tussen wel en niet opereren hebben wij bij gebrek aan gegevens niet de graad van ernst van de carotisstenose betrokken. In diverse publicaties kon echter geen verband worden gevonden tussen de prognose en de mate van vernauwing (14, 21, 40). Focale cerebrale ischémieën lijken eerder te worden veroorzaakt door embolieën dan door een mogelijke hemodynamische consequentie van een aanwezige stenose (41). Bovendien, ook al zou het potentiële voordeel van operatie toenemen bij ernstiger stenosen, de risico's van angiografie en operatie worden dan waarschijnlijk ook groter.

Over patiënten met bilaterale stenosen waren te weinig specifieke gegevens voorhanden voor een aparte analyse. In het experiment met symptomatische patiënten leek weliswaar de prognose voor deze groep patiënten te verbeteren na geslaagde operatie, maar dit werd overschadwd door een overmaat aan operatiecomplicaties (11).

Volgens sommigen is operatie effectief als zij wordt gereserveerd voor patiënten jonger dan 65 jaar zonder risicofactoren als hypertensie, hartinfarct of decompensatio cordis in de anamnese (6, 23). Dit wordt echter niet ondersteund met empirische gegevens. Het feit dat zulke patiënten er na operatie naar verhouding goed aan toe zijn, is bedrieglijk. Bij gebrek aan een geschikte controlegroep kan het effect van operatie worden verward met een reeds a priori betere prognose.

Bij de beoordeling van de effectiviteit van endarterectomie zijn ook de volgende overwegingen van belang:

- Patiënten met carotisstenosen hebben dikwijls een meer algemeen vaatlijden (42, 43). Ook in verband met hun leeftijd zal er relatief vaak sprake zijn van aandoeningen als hypertensie, hartlijden en diabetes. In veel studies komt de meeste sterfte dan ook op rekening van cardiale pathologie (15, 23, 44, 45).
- De localisatie van een souffle of stenose heeft weinig voorspellende waarde voor de localisatie van later optredende accidenten (3, 42, 45).  
Om deze redenen ligt het voor de hand dat een locale ingreep maar een beperkt effect kan hebben, dat bovendien gemakkelijk te niet wordt gedaan door de risico's van operatie.
- Tenslotte kan er re-stenoserig optreden, tot in 3.5% van de gevallen per jaar na operatie (46-52).

Onze analyse was gericht op angiografisch aangetoonde stenosen. Recent onderzoek met behulp van non-invasieve methoden (Duplex Doppler scan) ondersteunt de opvatting dat een conservatieve benadering van patiënten met een asymptomatische carotissouffle de voorkeur verdient (53, 54).

#### Conclusie.

Op grond van de literatuur kan operatieve behandeling van asymptomatische carotisstenosen niet gerechtvaardigd worden. Wij menen dat endarterectomie niet als behandeling van asymptomatische carotisstenose aanvaard kan worden als de effectiviteit ervan niet is aangetoond.

#### LITERATUUR.

1. Lodder J, De Krom MCTFM. De consequenties van een toevalsbevinding: de asymptomatische carotissouffle. Ned T Geneesk 1982; 126: 1220-1225.
2. Thompson JB, Patman RD, Persson AV. Management of asymptomatic carotid bruits. American Surgeon 1978; 42: 77-80.
3. Hurst JW. Noises in the neck. N Engl J Med 1980; 302: 862-863.
4. Strandness DE. Why the controversy about carotid endarterectomy. J Cardiol Med 1981; 6: 1067-1070.

5. Barnett HJM, Plum F, Walton JN. Carotid endarterectomy - an expression of concern. *Stroke* 1984; 15: 941-943.
6. Thompson JE, Patman RD, Talkington CM. Asymptomatic carotid bruit: longterm outcome of patients having endarterectomy compared with unoperated controls. *Ann Surg* 1978; 188: 308-316.
7. Busuttil RW, Baker JD, Davidson RK, Machleder HI. Carotid artery stenosis - hemodynamic significance and clinical course. *JAMA* 1981; 245: 1438-1441.
8. Fields WS. The asymptomatic carotid bruit - operate or not? *Stroke* 1978; 9: 269-271.
9. Mohr JP. Asymptomatic carotid artery disease. *Stroke* 1982; 13: 431-433.
10. Kuller LH, Sutton KC. Carotid Artery Bruit: is it safe and effective to auscultate the neck. *Stroke* 1984; 15: 944-947.
11. Fields WS, Maslenikov V, Meyer JS, Hass WK, Remington RD, MacDonald M. Joint study of extracranial arterial occlusion. *JAMA* 1970; 211: 1993-2003.
12. Lubsen J, Hunink MGM. Medische besliskunde: een oud probleem in een nieuwe jas. *Ned Tijdschr Geneesk* 1984; 128: 249-257.
13. Hennerici M, Rautenberg W, Mohr S. Stroke risk from symptomless extracranial arterial disease. *Lancet* 1982; 1180-1183.
14. Kroener JM, Dorn PL, Schoor PM, Wickbom I, Bernstein EF. Progress of asymptomatic ulcerating carotid lesions. *Arch Surg* 1980; 115: 1387-1392.
15. Levin SM, Sondheim FK, Levin JM. The contralateral diseased but asymptomatic carotid artery: to operate or not? An update. *Am J Surg* 1980; 140: 203-205.
16. Moore WS. Natural history of nonstenotic, asymptomatic ulcerative lesions of the carotid artery. *Arch Surg* 1978; 113: 1352-1359.
17. Moore WS, Boren C, Malone JM, Goldstone J. Asymptomatic carotid stenosis: immediate and long term results after prophylactic endarterectomy. *Am J Surg* 1979; 138: 228-233.
18. Weinstein MC, Fineberg HV. *Clinical Decision Analysis*. Saunders, Philadelphia 1980.
19. Pauker SG, Kassirer JP. *Decision Maker. The personal micro-computer decision analysis*, 1982; version 3.0. New England Medical Center, Boston.

20. Humphries AW, Young JR, Sentilli PH, Beven EG, de Wolfe VG. Unoperated, asymptomatic significant internal carotid artery stenosis: A review of 182 instances. *Surgery* 1976; 80: 695-698.
21. Podore PC, De Weese JS, May AG, Rob CG. Asymptomatic contralateral carotid artery stenosis: a five years follow-up study following carotid endarterectomy. *Surgery* 1980; 88: 748-752.
22. Durward QJ, Ferguson GG, Barr HWK. The natural history of asymptomatic carotid bifurcation plaques. *Stroke* 1982; 13: 459-464.
23. Javid HJ, Ostermiller WE, Hengesh JW, et al. Carotid endarterectomy for asymptomatic patients. *Arch Surg* 1971; 102: 389-391.
24. Nunn DB. Carotid endarterectomy: an analysis of 234 operative cases. *Ann Surg* 1975; 182: 733-738.
25. Easton JD, Sherman DG. Stroke and mortality rate in carotid endarterectomy: 228 consecutive operations. *Stroke* 1977; 8: 565-568.
26. Sacco RL, Wolf PA, Kannel WB, McNamara PM. Survival and recurrence following stroke, The Framingham Study. *Stroke* 1982; 13: 290-295.
27. Solzi P, Ring H, Najenson T, Luz Y. Hemiplegics after a first stroke: late survival and risk factors. *Stroke* 1983; 14: 703-709.
28. Garaway WM, Whisnant JP, Drury I. The changing pattern of survival following stroke. *Stroke* 1983; 14: 699-703.
29. Toole JF, Yuson CP, Janeway R, et al. Transient Ischemic Attacks: A prospective study of 225 patients. *Neurology* 1978; 28: 746-753.
30. Whisnant JP, Cartlidge NEF, Elveback LR. Carotid and vertebral-basilar transient ischemic attacks; effect of anti-coagulants, hypertension and cardiac disorders on survival and stroke occurrence - a population study. *Annals of Neurology* 1978; 3: 107-115.
31. Heyden S, Heiss G, Heyman A, et al. Cardiovascular Mortality in Transient Ischemic Attacks. *Stroke* 1980; 11: 252-255.
32. Simonsen N, Christiansen D, Heltberg A, Marquardsen J, Pedersen E, Sorensen PS. Long-term prognosis after transient ischemic attacks. *Acta Neurol Scandinav* 1981; 63: 156-168.
33. Cote R, Barnett HJM, Taylor DW. Internal carotid occlusions: a prospective study. *Stroke* 1983; 14: 898-902.
34. Eisenberg RL, Bank WO, Hedcock MW. Neurologic complications of angiography for cerebrovascular disease. *Neurology* 1980; 30: 895-897.
35. West H, Burton R, Roon AJ, Malone JM, Goldstone J, Moore WS. Comparative risk of operation and expectant management for carotid artery disease. *Stroke* 1979; 10: 117-121.

36. Hass JS. An approach to the maximal acceptable complication rate after surgery for Transient Cerebral Ischemia (TIA). *Stroke* 1979; 10: 84.
37. Thompson JE, Talkington CM. Carotid endarterectomy. *Ann Surg* 1976; 184: 1-15.
38. Thompson JE. Complications of carotid endarterectomy and their prevention. *World J Surg* 1979; 3: 155-163.
39. Brott T, Thalinger K. The practice of carotid endarterectomy in a large Metropolitan Area. *Stroke* 1984; 15: 950-955.
40. Johnson N, Burnham SJ, Flanigan DP, Goodreau JJ, Yao JST, Bergan JJ. Carotid endarterectomy: a follow-up study of the contra-lateral non-operated carotid artery. *Ann Surg* 1978; 188: 748-752.
41. Fell G, Breslau P, Knox RA, Philips D, Thiele BL, Strandness DE. Importance of non-invasive ultrasonic Doppler testing in the evaluation of patients with asymptomatic carotid bruits. *Am Heart J* 1981; 102: 21-26.
42. Heyman A, Wilkinson WE, Heyden S, et al. Risk of stroke in asymptomatic persons with cervical arterial bruits: a population study in Evans County, Georgia. *N Engl J Med* 1980; 302: 838-841.
43. Levin SM, Sondeimer FK. Stenosis of the contralateral carotid artery - to operate or not? *Vasc Surg* 1973; 7: 3-13.
44. Kartchner MM, McRae PL. Non-invasive evaluation and management of the "asymptomatic" carotid bruit. *Surgery* 1977; 82: 840-847.
45. Wolf PA, Kannel WB, Sorlie P, McNamara P. Asymptomatic carotid bruit and risk of stroke. The Framingham Study. *JAMA* 1981; 245: 1442-1445.
46. Callow AD, Matsumoto G, Cossman D, Stein A. Early restenosis after carotid endarterectomy. *Stroke* 1977; 8: 14.
47. Cantelmo NL, Culter BS, Wheeler HB, Herrmann JB, Cardullo PA. Non-invasive detection of carotid stenosis following endarterectomy. *Arch Surg* 1981; 116: 1005-1008.
48. Cossman D, Callow AD, Stein A, Matsumoto G. Early restenosis after carotid endarterectomy. *Arch Surg* 1978; 113: 275-278.
49. Hertzner NR, Martinez BD, Benjamin SP, Beven EG. Recurrent stenosis after carotid endarterectomy. *Surg Gyn Obstet* 1979; 149: 360-364.
50. Kremen JE, Kaupp HA, McDonald KM. Restenosis or occlusion after carotid endarterectomy. *Arch Surg* 1979; 114: 608-610.

51. Ortega G, Gee W, Kaupp HA, McDonald KM. Postendarterectomy carotid occlusion. Surgery 1981; 1093-1097.
52. Stoney RJ, String ST. Recurrent carotid stenosis. Surgery 1976; 80: 705-710.
53. Lees RS. The natural history of carotid artery disease. Stroke 1984; 15: 603-604.
54. Roederer GO, Langlois YE, Jager KA. The natural history of carotid arterial disease in asymptomatic patients with cervical bruits. Stroke 1984; 15: 605-613.

#### APPENDIX

Leeftijdscorrectie voor de sterftekans.

1. Omdat de gemiddelde leeftijd in de conservatief behandelde patiëntengroepen 63 jaar was en die in de geopereerde groepen 66 jaar, pasten wij een leeftijdscorrectie toe op de sterftekans van de eerste groep.

De 4-jaars sterftekans voor 63-jarigen is 0.256.

De 4-jaars sterftekans in de algemene bevolking voor 63-jarigen is voor mannen en vrouwen tezamen: 0.070.

Het extra sterfterisico ("excess-risk") voor deze patiënten is dan:  $0.256 - 0.070 = 0.186$ .

De 4-jaars sterftekans in de algemene bevolking van 66-jarigen is voor mannen en vrouwen tezamen: 0.093.

Als de (voor de leeftijd gecorrigeerde) 4-jaars-sterftekans van conservatief behandelde asymptomatische patiënten namen wij nu: "excess-risk" + leeftijdsspecifieke 4-jaars-sterftekans =  $0.186 + 0.093 = 0.279$ .

2. Op dezelfde wijze corrigeerden wij voor leeftijdsverschillen de sterftekans van:

- de conservatief en operatief behandelde patiënten in de tweede 4-jaars-periode. Zij zijn immers gemiddeld 4 jaar ouder.

- de sterftekans na het optreden van CVA en TIA. De gemiddelde leeftijd van deze patiënten in de gepubliceerde series was namelijk 67 respectievelijk 60 jaar, terwijl in de analyse bij het begin van de tweede 4-jaars-periode uitgegaan wordt van een leeftijd van 70 (= 66 + 4) jaar.

## Hoofdstuk 19

# Over vertekening van testuitslagen door reeds aanwezige informatie

J.A. KNOTTNERUS EN Prof. Dr. F. STURMANS \*

**D**AGELIJKS worden er door huisartsen en specialisten onderzoeken aangevraagd bij 'derden'. Zoals bij alle diagnostische verichtingen is hierbij de bedoeling, dat vooraf ontwikkelde hypothesen, gevormd op grond van anamnese en reeds verricht onderzoek, door de uitslag van de aangevraagde tests op hun waarschijnlijkheid getoetst worden. Anders gezegd: door de (interpretatie van de) testuitslag veranderen de 'a priori kansen' op bepaalde aandoeningen of toestanden in 'a posteriori kansen'

**Samenvatting** *Bij het aanvragen van diagnostisch onderzoek wordt dikwijls een korte beschrijving gegeven van het klinisch beeld. Vaak is er ook een gerichte vraagstelling. De door de diagnost afgegeven testuitslag kan dan rechtstreeks beïnvloed worden door de vooraf verkregen informatie. Het onderzoek is dan geen onafhankelijke toetssteen meer van de diagnostische hypothesen van de aanvragende arts.*

(fig. 1). De grootte van die veranderingen zijn bepalend voor de mate waarin de testuitslag zou mogen bijdragen aan de besluitvorming.

### **Betekenis van reeds aanwezige informatie ('prior' informatie)**

Er zijn verschillende procedures gangbaar wat betreft de aan de te consulteren diagnost te verstrekken informatie. Bij bloedonderzoek worden in het algemeen slechts de naam, de geboortedatum, het adres en het ziekenfondsnummer op het aanvraagformulier genoteerd. Bij andere onderzoeken, zoals röntgendiagnostiek, elektrocardiografie, endoscopie, pathologisch-anatomisch onderzoek en dikwijls ook bij microbiologisch onderzoek ligt dit anders. In de regel wordt een korte beschrijving gegeven van het klinische beeld, gekoppeld aan een differentieële diagnose en/of een gerichte vraagstelling.

Bij bloedonderzoek dat elders verricht wordt, zal het niet vaak voorkomen dat de

uitvoering ervan wordt beïnvloed door reeds aanwezige informatie, ook al, omdat de aard van de bepalingen zich daarvoor niet leent. In de andere genoemde gevallen echter is te verwachten dat de diagnost terdege met de vooraf verstrekte informatie rekening houdt.

Dit heeft enerzijds voordelen: de diagnost kan zijn aandacht richten op de vraagstelling, en speciaal letten op aanwijzingen in die richting. Soms ook zal hij ermee rekening houden in verband met de specifieke eisen die gesteld moeten worden aan de te volgen procedures, bij voorbeeld als er bij microbiologisch onderzoek een speciale voedingsbodem wenselijk is. Ten slotte zal de diagnost, bij het presenteren van de uitslag, het door hem gevonden beeld kunnen interpreteren in het licht van de vooraf reeds bekende gegevens.

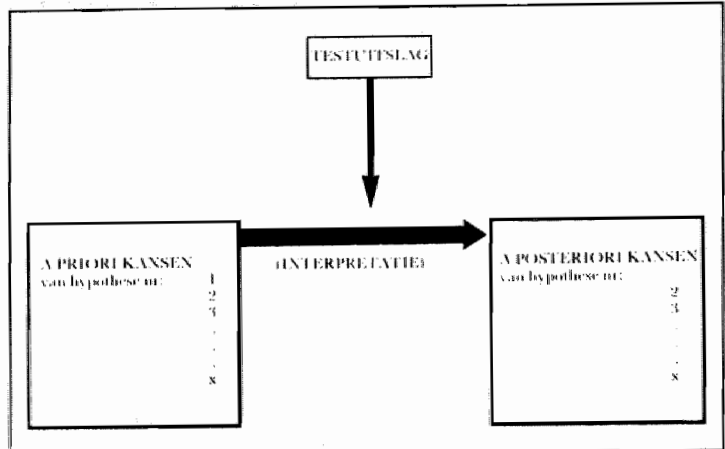
\*Arts, resp. hoogleraar epidemiologie  
Capaciteitsgroep Epidemiologie/Gezondheidszorgonderzoek  
Rijksuniversiteit Limburg te Maastricht



Anderzijds schuilt er een ad-der onder het gras: namelijk dat onbewust de gerichtheid op een bepaald punt omslaat in kokerzien dat testuitslagen op zichzelf niet onafhankelijk van de vooraf reeds bekende informaties en interpretaties tot stand komen, maar hierdoor in sterke mate beïnvloed en mogelijk vertekend worden. Dat wil zeggen, de testuitslag zelf wordt mede bepaald door de a priori kans; daarnaast wordt bij de interpretatie van de uitslag de a priori kans ook weer (en in dit geval terecht) in beschouwing genomen, zodat de a priori informatie 'dubbel' wordt geteld. De test is dan geen echte test, in de zin van een onafhankelijke toetssteen, maar voor een deel slechts een middel tot 'bevestiging' van vermoedens (fig. 2). Het gevolg kan zijn dat nu te gemakkelijk een bepaalde diagnose definitief aanvaard wordt, of juist te weinig aandacht krijgt.

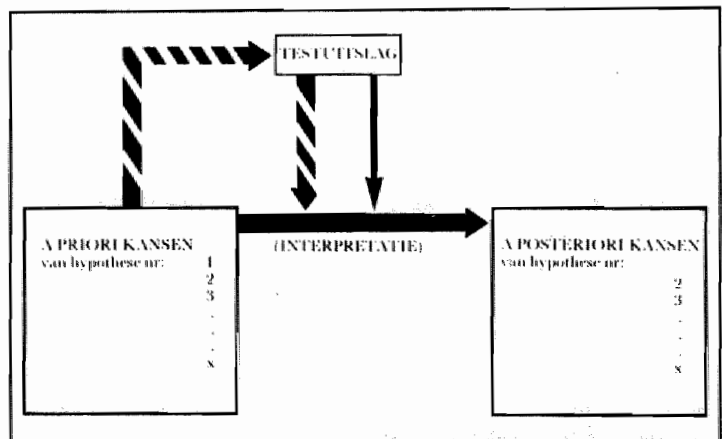
### Praktische problemen

Als men het principe aanvaardt dat testuitslag en a priori informatie van elkaar onafhankelijk moeten zijn, dan blijkt direct dat dit in de praktijk niet altijd is vast te houden. Het meest duidelijk is dit, als hypothesevorming en het uitvoeren van tests door één



Figuur 1.

Door interpretatie van de testuitslag gaan de a priori kansen over in a posteriori kansen.



Figuur 2.

De testuitslag wordt beïnvloed door de a priori kansen. Bij interpretatie van de testuitslag in het licht van de a priori kansen wordt de a priori informatie 'dubbel' geteld.

en dezelfde persoon plaatsvinden. De arts die op grond van de anamnese overmatig alcoholgebruik vermoedt, zal vervolgens wellicht de lever eerder palpabel achten dan wanneer een dergelijke anamnese zou ontbreken. Ook is denkbaar dat de bloeddrukmeting bij een zwaar rokende, ongezond levende manager naar boven wordt afgerond; terwijl bij een sportieve patiënt van wie men weet dat deze een hekel heeft aan medicijnen, het tegenovergestelde kan gebeuren. En denkt u eens aan het aflezen van de kleurnuances van de Sicca-meter bij iemand met onverklaarde moeheidsklachten. In dergelijke gevallen zijn hypothesevorming en diagnostiek moeilijk van elkaar te scheiden. Men kan dan niet veel meer doen dan:

1. toch proberen zo objectief mogelijk te werk te gaan;
2. er rekening mee houden dat de bevindingen beïnvloed kunnen zijn door de voorkennis.

Een arts die met een collega samenwerkt kan, als het om voor het beleid cruciale bevindingen gaat, deze vragen het onderzoek ook bij dezelfde patiënt te doen zonder voorafgaande anamnese. De a priori informatie dient dan ook niet op andere subtiele wijzen te worden overgedragen. Beter

dan: 'Zou jij ook even willen voelen of je de lever van meneer A ook vergroot vindt?', is: 'Zou jij de buik van meneer A even willen onderzoeken?' Een systematisch buikonderzoek is dan wel een vereiste.

Terugkerend naar de situatie waarbij onderzoek bij derden wordt aangevraagd, lijkt het daar allemaal op het eerste gezicht minder moeilijk te liggen. En bij bloedonderzoek is dat ook zo. Maar in andere gevallen, zoals radiodiagnostiek of pathologisch-anatomisch onderzoek, kan de genoemde 'dubbel telling' juist extra sterk uitpakken:

- De aanvragende arts verstrekt a priori informatie.
- De diagnost beoordeelt en beschrijft de uitslag in het licht van de a priori informatie.
- De aanvragende arts meent, na ontvangen van de uitslag, dat nu eindelijk objectieve gegevens bekend zijn, en op zijn beurt interpreteert hij de beschreven uitslag in het licht van de a priori informatie.

Het moge duidelijk zijn dat op deze wijze de 'voorspellende' waarden van bij voorbeeld een uitstrijkje of een mammogram hoger kunnen uitvallen dan zij zouden zijn als de reeds bekende informatie niet in de

uitslag verwerkt was. Als men hierbij voorts het vaak kleurrijke, maar niet altijd voor één interpretatie vatbare taalgebruik van veel diagnosten in aanmerking neemt, worden de volgende uitslagen begrijpelijk: 'Inderdaad wat grove maagplooiën, zoals men die wel ziet bij chronische gastritis'; 'Op zichzelf geen verdacht beeld, doch maligniteit niet geheel uit te sluiten'; 'Enige discussversmalling op de niveaus L3-L4-L5. Gezien de anamnese valt myelografie te overwegen'.

### **Een voorbeeld: de biopsie**

Het probleem van veel uitslagen is niet zozeer het feit dat geen volstrekt zekere uitspraak gedaan wordt. Immers, waar zekerheid niet is te geven, dient deze ook niet gesuggereerd te worden. De moeilijkheid is vooral dat de mate van onzekerheid vaak onduidelijk blijft. Dit geldt a fortiori als het gaat om uitslagen, waarbij de grens tussen normaal en abnormaal niet scherp te trekken is en als er meer dan één diagnostisch alternatief overwogen moet worden. Het zijn ook juist die situaties, waarin de voorkennis een op zichzelf niet typisch beeld het laatste duwtje kan geven in de ene of de andere richting. Ook ten aanzien van

pathologisch-anatomisch onderzoek, dat vaak geacht wordt de definitieve arbitrage – de 'gouden standaard' – te zijn, zijn er wat dit betreft meer dilemma's dan in het algemeen vermoed wordt. Underwood geeft in zijn 'Introduction to biopsy interpretation and surgical pathology' het volgende voorbeeld: een botbiopsie wordt genomen van een laesie, waarvan op klinische en radiologische gronden wordt vermoed dat het een osteosarcoom is. De biopsie laat in werkelijkheid niets anders zien dan een fractuurcallus. Een patholoog geeft als uitslag: osteosarcoom. Hij maakt deze beoordelingsfout op grond van de a priori informatie en omdat hij weet dat callusweefsel sterk op osteosarcoom kan lijken. Gezien de hem reeds verstrekte informatie neemt hij het zekere voor het onzekere. Hij moet immers van goeden huize komen om de clinicus en de radioloog tot andere gedachten te brengen, en het zal toch maar eens een osteosarcoom zijn. Een andere patholoog rapporteert wel een callusbeeld, maar men acht hem weinig geloofwaardig omdat hij bekend staat als iemand die uitslagen afgeeft zonder het röntgenbeeld gezien te hebben. Een derde patholoog, die op zich wel ge-

loofwaardig wordt geacht, herkent de callusvorming ook. Deze conclusie is echter niet in overeenstemming met de klinische en radiologische bevindingen. De patholoog gaat er vanuit dat dit slechts geaccepteerd wordt als dit gebrek aan overeenstemming expliciet wordt vermeld, met de mogelijke verklaring dat dit kan liggen aan een inadequaat gekozen plek om te biopteren.<sup>1</sup>

Underwood beschrijft voorts aan de hand van vele voorbeelden de geleidelijke overgang van 'normaal' naar maligne, en van benigne naar maligne. Vaak zijn goed- en kwaadaardig niet precies te onderscheiden. Er zijn veel geheel of gedeeltelijk overlappende morfologische kenmerken, ook als het om andere laesies dan nieuwvormingen gaat. Weinig beelden zijn pathognomonisch en foute classificaties zijn zeer wel mogelijk. Wat betreft de vermelding van uitslagen aan de aanvragers stelt Underwood dat het zinvol is niet alleen te melden dat het beeld past bij bepaalde afwijkingen, maar ook dat het bij andere afwijkingen niet past. Dit kan helpen relevante diagnoses uit te sluiten. Wat betreft het komen tot meer eensluidende, kwantificeerbare uitslagen, wijst hij in

navolging van anderen op de mogelijkheid om diagnoses te vermelden met een bepaald betrouwbaarheidsinterval. Hierdoor zou de mate van onzekerheid kunnen worden overgedragen.

### De methode-Schwartz

Een meer precieze, besliskundige benadering is die van Schwartz et al.<sup>2</sup> Zij gaan er daarbij van uit dat diagnostiek, ook de pathologisch-anatomische, gekenmerkt wordt door het werken met onzekerheden en dus kansen. Met hun methode wordt tegemoetgekomen aan de behoefte aan zowel kwantificering van onzekerheden, als aan de noodzaak van het onafhankelijk zijn van de testuitslag van de a priori informatie. De methode is als volgt:

1. de behandelende arts stelt op grond van het klinische beeld een lijst op van aandoeningen en toestanden die in aanmerking komen. Voor zichzelf geeft hij door middel van een numerieke schatting aan hoe waarschijnlijk de alternatieven zijn ten opzichte van elkaar (a priori kansen);
2. de patholoog ontvangt bij wijze van vraagstelling de lijst van alternatieven in willekeurige volgorde, zonder aanduiding van hun

TABEL 1.

Voorbeeld: het veranderen van de a priori kansen in de a posteriori kansen op diverse alternatieve diagnostische hypothesen onder invloed van de informatie van de biopsie (methode-Schwartz).

Diagnostische hypothese	A PRIORI, O.G.V. KLINISCH BEELD	TESTINFORMATIE	A POSTERIORI	
	(1) Relatieve waarschijnlijk- heid van elk alternatief vóór biopsie	(2) Relatieve waarschijnlijk- heid van het aangetroffen biopsiebeeld, uitgaand van elk alternatief	(1) x (2) Relatieve waarschijnlijk- heid van elk alternatief na biopsie	(1) x (2)/Σ. Kans op elk alternatief, na biopsie
(Variant van) normaal	15	1	15	18%
Ontsteking	15	2	30	35%
Benigne tumor	5	5	25	29%
Maligne tumor	1	15	15	18%
			Σ = 85	

respectieve waarschijnlijk-  
heden. Hij bekijkt het pre-  
paraat en geeft nu zelf in  
een numerieke schatting  
weer hoe waarschijnlijk het  
gevonden morfologische  
beeld is bij elk van de door  
de aanvrager genoemde al-  
ternatieven ('conditionele'  
kansen). De patholoog kan  
hierbij uiteraard ook alter-  
natieven betrekken die  
door de aanvragende arts  
niet waren overwogen;

- als de patholoog nieuwe  
mogelijkheden heeft ge-  
noemd, wordt de aanvrager  
eerst gevraagd deze als-  
nog te voorzien van een a  
priori kans, in relatie tot de  
andere alternatieven;
- ten slotte worden de schat-  
tingen van de patholoog,  
op grond van alleen de bi-  
opsie, gemeld aan de aan-  
vragende arts. Deze kan ze  
nu combineren met de ei-  
gen a priori schattingen, en  
hieruit volgt de a posteriori  
schatting. Dit gebeurt, for-  
meel gesproken, door de a  
priori waarschijnlijkheden  
van elk alternatief met de  
betreffende conditionele  
kansen te vermenigvuldi-

gen (appendix). In tabel 1  
is deze benadering aange-  
geven met een voorbeeld.  
In principe is zij dezelfde  
als die, welke aan de orde  
was in het eerste artikel in  
deze rubriek.<sup>3</sup> Het verschil  
is dat hier uitgegaan wordt  
van relatieve waarschijn-  
lijkheden in plaats van in  
een percentage uitgedrukte  
kansen. Het gaat dan om  
de waarschijnlijkheid van  
respectievelijk:

- Een bepaalde diagnose  
of toestand op grond van  
de informatie die voor het  
doen van de test over de  
betreffende patiënt be-  
schikbaar is (kolom 1).
- Het aangetroffen beeld,  
gegeven elk der overwogen  
alternatieven (kolom 2).
- Elk alternatief, na com-  
bineren van de informatie  
van de kolommen 1 en 2.  
Dit is weergegeven in de  
kolommen onder 'a posteri-  
ori', voor respectievelijk de  
relatieve waarschijnlijkhe-  
den en de in een percentage  
uitgedrukte kansen.

De relatieve waarschijnlijkhe-  
den zijn telkens uitgedrukt als

veelvouden van de minst  
waarschijnlijke mogelijkheid.  
Deze uitdrukkingsvorm heeft  
praktische voordelen als er  
veel alternatieven overwogen  
worden of wanneer sommige  
alternatieven een zeer lage  
waarschijnlijkheid hebben.  
Uit dit voorbeeld blijkt dat,  
als men de biopsie-informatie  
combineert met de a priori  
schattingen, het mogelijk is  
dat maligniteit uiteindelijk  
toch niet het meest waar-  
schijnlijk is. Dit ondanks het  
feit dat men dit biopsiebeeld  
het meest frequent (doch niet  
altijd!) aantreft bij maligne  
tumoren.

Schwartz c.s. vergeleek voor  
een viertal reële casussen met  
elkaar enerzijds de a posteriori  
kans op de diverse alternatie-  
ven volgens de zojuist be-  
schreven formele methode tot  
stand gekomen, met ander-  
zijds de rechtstreekse, meer  
intuïtieve, schatting van de a  
posteriori kansen door de pa-  
tholoog, op grond van zowel  
de klinische als de histologi-  
sche gegevens. Hierbij bleken  
grote verschillen naar voren te  
komen. Zo leidde in één geval  
de formele methode tot een

kans op tuberculose versus die op sarcoidose van 80%:20%, terwijl de directe interpretatie door de patholoog de verhouding 15%:85% opleverde.

### Conclusies

Het is niet altijd mogelijk, en zeker niet gebruikelijk, om de uitslag van een diagnostisch onderzoek te beschrijven los van de reeds aanwezige klinische vermoedens. Bovendien is men gewend om vooral naar één uitslag toe te schrijven, zonder explicitering van de kansen dat het gevondene voorkomt bij verschillende alternatieve diagnoses.

Een benadering, zoals die van Schwartz c.s., kan echter voordelen bieden voor de rapportage en interpretatie van uitslagen van diverse diagnostische technieken.

De methode houdt namelijk in:

1. een duidelijker scheiding van beschrijving en interpretatie van de uitslag;
2. geringere afhankelijkheid van de uitslag van reeds bestaande vermoedens;
3. kwantificering van onzekerheden, en daarmee het wegnemen van veel spraakverwarring;
4. het mogelijk maken van beslistkundig-analytische benaderingen met voordelen voor onderwijs, evaluatie

### Appendix

In de benadering van Schwartz c.s. wordt impliciet gewerkt met het begrip 'aannemelijkheidsquotiënt' (in het Engels: likelihood ratio).

Dit is:

de kans op een bepaalde testuitslag bij mensen met een bepaalde ziekte

de kans op die uitslag bij gezonden

Uit kolom 2 van tabel 1 is dit quotiënt voor elk alternatief direct af te lezen, namelijk:	
voor ontsteking	$2/1 = 2$
voor benigne tumor	$5/1 = 5$
voor maligne tumor	$15/1 = 15$

Het aannemelijkheidsquotiënt is een maat voor het onderscheidend vermogen van de testuitslag. Als het gaat om een situatie waarin slechts sprake is van de mogelijkheden ziek en niet-ziek zijn, is het aannemelijkheidsquotiënt gelijk aan:

de sensitiviteit

100% - de specificiteit

- en toetsing;
5. een heldere afbakening van verantwoordelijkheden van diagnost en clinicus.

Als literatuurgegevens ontbreken voor het onderbouwen van kwantitatieve schattingen, is een zo goed mogelijk subjectieve schatting op grond van de eigen ervaring altijd nog beter dan een mistige en voor velerlei uitleg vatbare formulering.

Afspraken tussen curatief en diagnostisch werkzame artsen zijn inmiddels zinvol (juist ook voor degenen die van de

geschetste formele benadering vooralsnog geen heil verwachten) teneinde te komen tot:

- Niet voor misverstand vatbare communicatievormen.

- Duidelijkheid over de te verstrekken voorinformatie en over de vraag of de integratie van klinische en diagnostische gegevens reeds door de diagnost verricht moeten worden. Voorts is het met het oog op de gewenste onafhankelijkheid van klinische voorinformatie en testuitslag van belang dat diagnostische methoden die hierop inspringen, aandacht krijgen, zoals:

- Het gebruik van "random-zero" bloeddrukmeters.
- Het volgen van standaardprocedures bij het beoordelen van foto's en preparaten, los van de aanvraagindicatie.

Het is niet duidelijk waarom dergelijke 'geblindeerde' beoordelingen in het algemeen slechts toegepast zouden moe-

ten worden in het kader van wetenschappelijk onderzoek en niet in het kader van patiëntenzorg. In de gedachtenvorming over vroegtijdige diagnostiek van mammacarcinoom wordt mede in dit verband onafhankelijke beoordeling door meer dan één radioloog resp. patholoog-anatoom overwogen. □

## Literatuur

- 1 Underwood, J.C.E.: *Introduction to biopsy interpretation and surgical pathology*. Springer-Verlag; Berlin, Heidelberg, New York, 1981.
- 2 Schwartz, W.B.; Wolfe, H.J.; Pauker, S.G.: *Pathology and probabilities, a new approach to interpreting and reporting biopsies*. New Engl. J. Med., 1981, 305: 917-923.
- 3 Knottnerus, J.A.; Sturmans, F.: *Het belang van de niet-afwijkende bevinding voor de differentiele diagnostiek*. The Practitioner, 1984, 1: 147-151.



## DEEL IV

### DIAGNOSTISCHE TESTS NADER BEZIEN.



DIAGNOSTISCHE TESTS.

Normen, referentiewaarden en testkenmerken nader bezien: een inventarisatie.

J.A. Knottnerus 1)

Samenvatting.

In het voorgaande is wat betreft de waarde van diagnostische tests meestal gebruik gemaakt van de termen sensitiviteit en specificiteit. Deze impliceren een bepaalde uitwerking van het begrip 'referentiewaarde'. In het nu volgende hoofdstuk wordt dieper op dit begrip ingegaan. Er wordt een inventarisatie gegeven van de verschillende wijzen waarop het vraagstuk van de referentiewaarden kan worden benaderd. Aan de orde komen de volgende benaderingen: gebruikelijk bij gezonden; onderscheiding van personen met en zonder bepaalde klachten of ziekten; testuitslag en prognose; testuitslag en het belang van de uitkomst. In dit kader worden diverse methoden en maten besproken. Het blijkt, dat referentiewaarden niet in het algemeen zijn te definieren, maar dat de te kiezen oplossing afhangt van de situatie en het gestelde doel.

1) Capaciteitsgroep Epidemiologie en Capaciteitsgroep Huisartsgeneeskunde

Rijksuniversiteit Limburg

## 1. Inleiding.

Resultaten van diagnostisch onderzoek ontleen hun betekenis aan het feit, dat zij iets zeggen over de gezondheidstoestand van mensen, en dat het bekend worden ervan nuttig is voor die gezondheidstoestand of, meer algemeen, het welzijn ten goede komt. In de praktijk komt het er op neer dat men bepaalde resultaten vergelijkt met een norm of referentiewaarde. Uit die vergelijking vloeit een conclusie voort die in de eerste plaats een interpretatie omvat van uitslagen als 'normaal' (of niet-afwijkend, negatief) of, in meer of mindere mate, afwijkend (positief). In de tweede plaats kan de testuitslag gevolgen impliceren voor de besluitvorming over eventuele verdere stappen op het gebied van diagnostiek en/of therapie.

Vanwege het normatieve aspect van het adjectief 'normaal', de mogelijke verwarring met de 'normale verdeling' volgens Gauss en tenslotte het besef dat er altijd een arbitrair element kleeft aan normen, spreekt men heden ten dage liever van referentiewaarden dan van normaalwaarden.

De interpretatie van een uitslag als vallend binnen of buiten het 'referentiegebied' kan op uiteenlopende wijzen tot stand komen, die niet tot dezelfde conclusie hoeven te leiden. Omdat een andere interpretatie van een uitslag kan leiden tot een ander therapeutisch beleid, is het van belang om bij de diverse benaderingen stil te staan. Vaak wordt er een onderscheid gemaakt tussen diverse typen parameters, dat voor onze bespreking van belang is, namelijk in:

- a) parameters die slechts aan- of afwezig kunnen zijn, of maar twee waarden kunnen aanemen (dichotome variabelen).
- b) parameters die meer dan twee, van elkaar onderscheiden, categorieën kennen. Hierbij kunnen weer worden onderscheiden:
  1. parameters waarbij aan de verschillende categorieën geen inherente volgorde ten grondslag ligt (nominale variabelen).
  2. parameters waarbij dit laatste wel het geval is (ordinale variabelen).
- c) parameters die op een bepaalde kwantitatieve schaal alle waarden kunnen aannemen (continue variabelen).

Voorbeelden van a) zijn: koliekpijnen, angina pectoris, een palpabele lever, het al dan niet aanwezig zijn van kanker; van b): haarkleur (nominaal), een cervix uitstrijk volgens Papanicolaou (ordinaal) en

van c): lengte, gewicht, bloeddruk en laboratorium parameters zoals het hemoglobinegehalte en de bezinking.

Men kan er over twisten in hoeverre genoemde onderscheiding opgaat. Bij als dichotoom en ordinaal opgevatte variabelen is bij nadere beschouwing toch vaak sprake van een glijdende schaal, waarbij op een bepaald moment een grens wordt overschreden (angina pectoris, een palpabele lever, een afwijkende glucose tolerantietest, overgangen tussen benigne en maligne histologisch beeld). Men kan ordinale en continue variabelen ook kunstmatig dichotomiseren, door één score of waarde te kiezen als grens tussen hoge en lage scores.

Met name het onderscheid tussen de typen a + b1 enerzijds en b2 + c anderzijds is voor onze bespreking van belang. Men kan bij de eerstgenoemde groep namelijk in de praktijk in het algemeen geen invloed uitoefenen op het positiviteitscriterium (de waarde van de testparameter die negatieve en positieve uitslagen scheidt). Bij ordinale en continue variabelen kan dit laatste wel. Er is immers een overgangsg gebied tussen hoge en lage waarden en men kan de referentie- of grenswaarden desgewenst manipuleren in één van beide richtingen.

Murphy onderscheidde zeven betekenissen van het begrip normaal: een 'normale' frequentieverdeling volgens Gauss volgend; representatief voor een groep; gebruikelijk; het meest gunstig ten aanzien van overleving en reproductie; zonder verhoogd risico op gezondheidsschade; bepaalde normen en conventies volgend; ideaal met voorbeeldfunctie (1). De laatste twee laten wij in het kader van de bespreking over diagnostische tests buiten beschouwing. In het vervolg hanteren wij, genummerd volgens de komende paragrafen, de indeling:

2. Normaal als gebruikelijk.

3. Relatie met klachten.

4. Relatie met ziekte.

4.1. Test en diagnose.

4.1.1. Sensitiviteit en specificiteit.

4.1.2. R.O.C. analyse.

4.1.3. Likelihoodratio en oddsratio.

4.1.4. Maximale toename van zekerheid.

4.2. Test en lange termijn prognose.

5. Relatie met het gewicht (de waarde) van de diverse uitkomsten.

5.1. Diagnostiek en behandelbaarheid.

5.2. Het optimale afkappunt.

5.3. De optimale teststrategie.

## 2. Normaal als gebruikelijk.

Aan deze benadering ligt in beginsel de volgende redenering ten grondslag: in geval van kenmerken of waarden die relatief weinig voorkomen is er sprake van een afwijking van de norm. Dextrocardie, roodharigheid of het hebben van bloedgroep AB zijn dan bijvoorbeeld 'afwijkingen'.

Deze benadering wordt veel toegepast bij laboratoriumparameters, die een continue verdeling volgen. In de rest van deze paragraaf zullen we hierop ingaan.

Men gaat vaak uit van de verdeling van waarden in een gezonde populatie en leidt hieruit af welke grenzen voor het referentiegebied dienen te gelden. Dikwijls postuleert men, al dan niet terecht, een 'normale' verdeling van de variabele of de logaritme ervan, en kiest als grenzen de waarden die 1.96 standaardafwijkingen van het gemiddelde liggen. Acht men een normale verdeling niet van toepassing, dan kan men een grens stellen bij een bepaalde percentiel. In het algemeen definieert men de 'normale' range zodanig dat deze de uitslagen van 95% van de personen in de populatie omvat. 'Normaal' is in deze benadering niets anders dan: gebruikelijk bij een gezonde (d.w.z. een niet als ziek te boek staande) populatie.

Een beperking van deze benadering is dat zij op zich zelf niet gerelateerd is aan de gezondheidstoestand van mensen, en daarmee geen aanknopingspunten biedt voor diagnostiek. Immers, de gevonden verdeling wordt niet afgezet tegen waarden bij zieken. Bovendien is er vaak discussie mogelijk over de vraag in hoeverre er sprake is van een "gezonde" populatie. Een voorbeeld: diverse onderzoekers pleiten voor het hanteren van andere normen voor het hemoglobinegehalte van blanken en negers op grond van de waarneming dat de hemoglobine-distributie voor negers in haar geheel 0.5 g/dl naar beneden verschoven is ten opzichte van blanken (2, 3). Micozzi meent dat de norm voor Aziatische zwangeren één g/dl lager mag worden gesteld dan voor Westerse zwangeren (4). Wintrobe, de nestor van de haematologie, ergert zich aan dergelijke benaderingen: "Locally prevalent values must not be equated with normal values. To do so would erroneously lead to the conclusion

Tabel 1. De prevalentie van Anemie in ontwikkelde landen in (subcategorieën van) de algemene populatie, zoals gevonden in diverse onderzoeken.

onderzoeker + jaar publicatie	populatie	steekproefgrootte		afrapport (%)		prevalentie (%)	
		m	v	m	v	m	v
Kilpatrick c.s. 1961	(6) algemene populatie Wales met follow-up door:	405	116	12,5	12	3	11,2
Watena c.s. 1969	(7) 35-64 jaar					2,2	6,9
Elwood c.s. 1967	(8) algemene populatie Wales, volwassen niet-zwangere vrouwen	1080		12			11
Garby c.s. 1967	(9) algemene populatie Uppsala vrouwen 15-46 jaar	5683		12			31
Hallberg c.s. 1968	(10) algemene populatie Göteborg vrouwen 35, 40, 45 jaar	289		12			24,6
Garby c.s. 1969	(11) algemene populatie Uppsala	5683		positieve respons* op ijzertherapie			17
Elwood c.s. 1972	(12) Algemene populatie Wales 65 jaar • Britse Aziaten • Engelsen • Welshmen	56 100 229	31 121 304	12 12 12		7 3 3,5	3 7 10
Bengtsson c.s. 1973	(13) algemene populatie Göteborg vrouwen 38, 46 jaar	803		12			6,6
Bengtsson c.s. 1978	(14) algemene populatie Göteborg vrouwen 44 jaar	326		12			7,4
Campbell c.s. 1981	(15) algemene populatie Gishorn (Nw. Zeeland) 65 jaar	196	361	12	12	4,1	14,3
de Wijn en Rusbach 1981	(16) scholieren 14-18 jaar a) hoge welstand b) lage welstand		14-15 jr: 12 16 jr: 12,5 17-18 jr: 13	11,5		a) 1 b) 3	1 3
Nijenhuis 1985	(17) werknemers 14-18 jaar (aanstellingskeuring)	1582	1159	13	12		12 17
v.d. Hoeven 1975	(18) mannelijke werknemers 42-59 jaar	4158		12,5			1,1

\* hierbij was geen algemeen afrapport geldig, maar als positief geldt een stijging van het hemoglobinegehalte na orale ijzertherapie.

that in a particular subpopulation or geographic region a different lower standard applies. This error has been made in surveys of large populations." Dit, zo meent hij, "is a disservice to patients and misleads physicians" (5). Dat niet alleen tussen westerse en niet-westerse landen verschillen in de hemoglobineverdeling bestaan, en dat het ook niet zo'n vaart loopt met het aanpassen van normen aan locale situaties blijkt uit tabel I. Zo vond men eind jaren '60 in Wales bij 11 % van de vrouwen waarden onder de 12 g/dl, terwijl in een Zweeds onderzoek 25 % onder dezelfde norm viel.

Wat betreft de bezinkingssnelheid van de erythrocyten (BSE) zijn er vergelijkbare controversen. Het International Committee for Standardization in Haematology beveelt expliciet aan om hierbij waarden binnen een range, die 90 of 95% van de locale gezonde bevolking omvat, als normaal te beschouwen (19). Daarbij moet nog rekening worden gehouden met de methode van bepaling omdat bijvoorbeeld de methode van Westergren bij hoge waarden van de BSE hogere uitslagen geeft dan de methode Wintrobe (20). Op grond van deze uitgangspunten stelden Böttiger (21) en Lewis (20) algemene richtlijnen vast, geslacht (vrouwen hebben een hogere BSE) en leeftijd (de BSE stijgt met de leeftijd (22)) in aanmerking nemend. Miller vond hun correctie voor de leeftijd nog onvoldoende - te veel ouderen zouden een te hoge bezinking krijgen toegedicht en stelde de 98<sup>e</sup> percentiel als bovengrens voor (23). Volgens anderen echter leidt correctie voor de leeftijd er toe dat bij ouderen een belangrijk signaal voor de ontdekking van pathologie wordt gemaskeerd (24, 25).

De uiteindelijke keuze van een grenswaarde kan zeer grote invloed hebben op het percentage afwijkende waarden. In tabel 2 hebben we aangegeven hoe vaak de BSE, zoals gemeten op het Diagnostisch Centrum Maastricht(DCM) in de jaren 1980 en 1982, als 'verhoogd' zou worden beschouwd bij toepassing van 3 verschillende normen: die van het Diagnostisch Centrum zelf, die volgens Lewis (de British Association of Clinical Pathologists (20)), en die welke recent door Miller zijn voorgesteld (23). De laatste twee benaderingen bevatten zoals gezegd een correctie voor geslacht en leeftijd, de eerste houdt alleen rekening met het geslacht. De verschillen zijn enorm: volgens de DCM-normen is 60 % verhoogd, volgens Lewis 40% en volgens Miller slechts 15 %.

Tabel 2: Percentage verhoogde BSE-waarden van op het Diagnostisch Centrum Maastricht onderzochte bloedmonsters, bij toepassing van drie verschillende normen.

	afwijkingen normen DC			afwijkingen normen Lewis 20)		afwijkingen normen Miller 23)	
	aantal	n	%	n	%	n	%
mannen	3149	1804	57,3	1270	40,3	471	15,0
vrouwen	1999	1296	64,8	770	38,5	301	15,1
totaal	5148	3100	60,2	2040	39,6	772	15,0

Het laatste voorbeeld geeft aan dat het hanteren van verschillende referentiewaarden tot grote spraakverwarringen kan leiden. Om dit te vermijden kan men voor een bepaalde populatie beter opgave doen van de gehele distributie dan van alleen het percentage waarnemingen onder (of boven) een bepaald afkappunt. Dat afkappunt is immers 'normatief' terwijl de gehele distributie daarvan niet afhankelijk is.

Onder het hoofdstuk 'normaal als gebruikelijk' kan tenslotte ook nog de 'individuele' norm worden ondergebracht. Als men uitgaat van een bepaalde gemeten uitgangswaarde voor een individu, kan men latere bepalingen in het licht hiervan interpreteren. Het kan voorkomen, dat iemand weliswaar enkele malen achtereen een uitslag heeft binnen het algemene referentie-gebied, maar dat tussen de afzonderlijke uitslagen grote verschillen bestaan: bijvoorbeeld een daling van het Hb bij een man van 10.5 mmol/l tot 8.6 mmol/l. Aan de andere kant kan iemand met een lichte teruggang schijnbaar 'abrupt' in het afwijkende gebied terecht komen (bijvoorbeeld bij een Hb-daling van 8.6 mmol/l tot 8.4 mmol/l). Bij de eerste patiënt is er echter meer reden tot ongerustheid dan bij de tweede patiënt. Costongs (26) demonstreerde hoe groot de intra-individuele variatie in de tijd kan zijn, ook bij gezonde proefpersonen. Dit kan weliswaar samenhangen met zich ontwikkelende pathologie, maar ook een gevolg zijn van fysiologische variabiliteit. Daarnaast moet men nog rekening houden met een eventueel beperkte reproduceerbaarheid van de test. Als maat voor het individuele referentiegebied stelde Costongs de "critical difference" voor: dit is de range waarbinnen, in een grotere populatie, 95% van de verschillen tussen twee aparte bepalingen van dezelfde parameter bij eenzelfde

individue vallen. Deze range is in het algemeen veel smaller dan de 95% range van de verdeling van de parameter voor alle personen gezamenlijk, op één moment in de tijd.

### 3. Relatie met klachten.

Klachten kan men in zekere zin 'tests' noemen, of soms ook 'ziekte' (als de klacht geheel of gedeeltelijk het gezondheidsprobleem zelf representeert, zoals bij hoofdpijn). Vanuit praktisch oogpunt echter is het zinvol apart aandacht te besteden aan het verband tussen subjectieve klachten en testuitslagen. De gedachtengang hierbij is als volgt: bepaalde klachten worden beschouwd als indicatie om een diagnostisch onderzoek te doen, dat dan weer bedoeld is om ziekte op het spoor te komen. Dit heeft meer zin, naarmate men in aanwezigheid van de klachten een duidelijker verschil in uitslagen vindt ten opzichte van klachtenvrije personen. Met andere woorden: er moet een verband zijn tussen de klacht en de testuitslag. Vindt men in beide gevallen hetzelfde of nagenoeg hetzelfde patroon, dan moet men overwegen de betreffende klacht als specifieke indicatie voor het doen van de test af te schaffen.

De ingang van de diagnostiek is de klacht, en niet de test. Het gaat dan ook niet zo zeer om de vraag hoe vaak de klacht voorkomt bij mensen met een 'positieve' respectievelijk 'negatieve' testuitslag. Van belang zijn:

- a) De kans dat bij personen met de klacht ( $S^+$ ) een bepaalde bevinding ( $T^+$ ) gedaan wordt, én
- b) Hoeveel groter deze kans is, dan bij mensen zonder de klacht.

In feite gaat het hier om de voorspellende waarde van de klacht ten aanzien van de diagnostische bevinding:  $P(T^+|S^+)$ , respectievelijk het quotiënt van de voorspellende waarden van de aanwezigheid respectievelijk de afwezigheid van de klacht:

$$\frac{P(T^+|S^+)}{P(T^+|S^-)}$$

Om de klacht van betekenis te doen zijn voor de indicatiestelling voor de test is nodig:

- dat het genoemde quotiënt duidelijk groter is dan 1, én



- dat de voorspellende waarde van de klacht ook op zich zelf van betekenis is, qua grootte.

Immers aan de eerste voorwaarde kan voldaan zijn, terwijl de voorspellende waarde toch laag blijft. Dit kan gebeuren als de betreffende bevinding zeer weinig voorkomt.

Het verschil met de benadering in de vorige paragraaf is dat nu niet (alleen) gekeken wordt naar de verdeling van uitslagen en bevindingen in een 'gezonde' populatie, maar dat een vergelijking wordt gemaakt van groepen, op grond van het al dan niet aanwezig zijn van bepaalde klachten.

Men moet hierbij uiteraard rekening houden met de selectiemechanismen die ten aanzien van de klacht, de bevinding en hun onderlinge verband kan zijn opgetreden, en men dient qua selectie vergelijkbare referentiegroepen te kiezen (27, 28) (zie hoofdstuk 3).

Voorbeelden van de bestudering van het verband tussen de aanwezigheid van klachten enerzijds en bevindingen anderzijds zijn de door ons ingestelde onderzoeken naar het verband tussen onverklaarde moeheid en het Hb (29) en tussen hoofdpijn en hoge bloeddruk (30). Het eerste onderzoek richtte zich op de spreekuurpopulatie van de huisarts, en het tweede vond plaats in de algemene bevolking. In beide gevallen was de voorspellende waarde van de bestudeerde klacht ten aanzien van de bestudeerde parameter erg laag, en waren de verbanden tussen klacht en bevinding afwezig respectievelijk zeer zwak. Deze onderzoeken worden elders in deze bundel beschreven. In een onderzoek bij zwangeren, beschreven in hoofdstuk 8, vonden wij wel enig onderscheidend vermogen van klachten ten aanzien van een laag Hb.

In dit verband kan ook gewezen worden op ons onderzoek naar het verband tussen 'onverklaarde moeheid' en het serumferritinegehalte in de huisartspraktijk (31). Tussen beide werd geen verband gevonden, en tussen het Hb- gehalte en het serumferritine evenmin. Bij zowel de personen met moeheidsklachten als de controlepersonen kwam in 11% van de gevallen een serumferritinegehalte voor van 12 µg/l of lager, dat volgens velen bewijzend zou zijn voor ijzergebrek. Bij een dergelijk hoge schatting van de prevalentie dringt zich de vraag op of screening, los van klachten en Hb-gehalte, geïndiceerd zou zijn. Bij het niet bestaan van een verband tussen klachten en diagnostische bevin-

dingen zijn dus twee elkaar op zich niet-tegensprekende conclusies mogelijk:

- er is géén indicatie voor de test aan te geven op grond van de klachten,
- screening lós van klachten is noodzakelijk als men eventuele achterliggende bevindingen op het spoor wil komen.

Of men het diagnostisch onderzoek dan niet of juist altijd toepast hangt onder andere af van:

- het belang van het op de hoogte zijn van de bevinding op zich, bijvoorbeeld met het oog op de prognose.
- de behandelbaarheid van de eventuele onderliggende pathologie,
- de geschiktheid van de test als screeningsinstrument (ondermeer het niet te belastend zijn van het onderzoek).

Tussen klachten en bevindingen kan een causale relatie zijn (bijvoorbeeld indien een laag Hb moeheid zou veroorzaken, en een hoge bloeddruk hoofdpijn). Dit hoeft echter niet het geval te zijn. Vaak gaat het om 'correlaties' ten gevolge van een gemeenschappelijke oorzaak. Bijvoorbeeld de bevinding van een Q-top op het ECG en de klacht pijn op de borst kunnen als gemeenschappelijke oorzaak hebben een myocardinfarct.

#### 4. Relatie met ziekte.

Uiteindelijk ligt de rechtvaardiging van diagnostisch onderzoek in de voorspelling van de aanwezigheid van aandoeningen, de prognose, en mate van de behandelbaarheid. In deze paragraaf hanteren wij de volgende tweedeling:

- 1) diagnostische betekenis van de test
- 2) prognostische betekenis van de test.

De behandelbaarheid en de relatie tot de 'waarde' van de diverse uitkomsten komen in de volgende paragraaf aan de orde.

##### 4.1. Test en diagnose.

Wat betreft de maten voor het verband tussen een testuitslag en de aanwezigheid ziekte ten behoeve van de diagnostiek zijn er diverse opties. We bespreken ze achtereenvolgens.

#### 4.1.1. Sensitiviteit en specificiteit.

Men kan de validiteit van de test uitdrukken in de sensitiviteit en specificiteit ten aanzien van de ziekte, en met behulp daarvan, geven een bepaalde prevalentie, de voorspellende waarde bepalen van een bepaalde testuitslag voor het al dan niet vóórkomen van ziekte. Dit is in hoofdstuk 2 afdoende belicht. Hierbij moet men er op bedacht zijn, dat sensitiviteit en specificiteit afhankelijk kunnen zijn van de kenmerken van de te onderzoeken populatie. In het kader van de verschillen tussen extra- en intramurale gezondheidszorg zijn belangrijke factoren hierbij:

- de mate van ontwikkeling van het ziektebeeld (hoofdstukken 2 en 21);
- de mate van selectie in relatie tot de uitslag vóórdát men in de te testen onderzoeksgroep terechtkomt (hoofdstuk 21).

Een voorbeeld van het laatste geeft de publicatie van Bezemer et al. (32). Zij definieerden als geschikte referentiepopulatie voor patiënten die verwezen waren naar de polikliniek of afdeling interne geneeskunde: álle in het recente verleden verwezen patiënten met uitzondering van de patiënten bij wie uiteindelijk diagnosen werden gesteld waarvan bekend is dat ze de test variabele kunnen beïnvloeden. Hierbij werd, onder andere voor de calciumspiegel, gevonden dat de waarden van de testvariabelen in de klinische referentiegroep vaak hoger uitvielen dan in de poliklinische referentiegroep. Blijkbaar is er sprake geweest van een selectieproces van perifeer naar ziekenhuis, waarbij personen met hogere uitslagen relatief vaak opgenomen werden, ook al was er geen sprake van diagnosen die met de betreffende uitslag in verband konden worden gebracht.

In tabel 3 presenteren wij, ter illustratie, gegevens over sensitiviteit en specificiteit van de BSE ten aanzien van verschillende aandoeningen, zoals in diverse onderzoeken vastgesteld. Er blijken tussen gerapporteerde series belangrijke verschillen te bestaan, afhankelijk van de populatie, de aandoening en het gehanteerde afkappunt. In het algemeen geldt dat sensitiviteit en specificiteit telkens verre van 100% zijn. Dat wil zeggen dat er een grote overlap is van waarden van zieken en niet-zieken, en een grote kans op fout-positieve en fout-negatieve uitslagen.

Tabel 3: Sensitiviteit en specificiteit (percentages) van de BSE-bepaling volgens tien onderzoeken.

Onderzoeker(s)	populatie	afkappunt mml/l h	ziekte	sensitiviteit	specificiteit
Pinchler/ Shanks (33)	algemeen v	9	algemeen	83	91
Boyd/Hofbrand (34)	ziekenhuis m + v 65 <sup>+</sup>	20	algemeen	a	71
Rafnsson et al (35)	algemeen v	30	artritis	20	90
			osteoartrose	15	91
			reumafactor+	30	89
Kenny et al. (36)	verwezen m + v	"low": 5m : 7v "intermediate": Lewis "high": Böttiger	afwijkingen plasma-eiwit	95	45
				83	78
Knuth et al. (37)	verwezen v	14	adnexitis	68 12	91 81
Peyman (38)	verwezen m + v	10 20	kanker	75 52	a) a)
Rai (39)	autopsie <sup>+</sup> m + v 64 <sup>+</sup>	10 20 30	kanker	87 74 66	10 26 38
Asbjörnsen (40)	verwezen m + v	volgens Böttiger	actieve / niet-actieve Hodgkin	70	84
Williams et al (41)	verwezen kinderen m + v	20	actieve / niet-actieve Hodgkin	90	43
Salter (42)	verwezen m + v	10 20	maag-darm- kanker	80 60	a) a)

a) op grond van de gepubliceerde gegevens niet te berekenen.

Een test kan voor verschillende aandoeningen een verschillende sensitiviteit en specificiteit hebben. Het is dan handiger niet te spreken van de sensitiviteit maar, neutraler, van de kans op een positieve uitslag bij ziekte-toestand  $D_x$  :

$$P(T^+ | D_x)$$

Wat betreft de specificiteit zou verwarring kunnen voorkomen tussen de kans op een negatieve uitslag bij personen zonder enige ziekte ( $P(T^- | D^-)$ ), en de kans op een negatieve uitslag bij personen zonder een bepaalde ziekte  $D$  ( $P(T^- | D_x^-)$ ). De tweede groep personen kan immers

wel degelijk zieken bevatten, namelijk die met andere aandoeningen dan  $D_x$ . Voor het rekenen met kansen ten aanzien van meer dan één toestand (waaronder gezondheid), kan men de respectievelijke kansen op positieve uitslagen bij de diverse toestanden (conditionele kansen) als uitgangspunt nemen. Gegeven een bepaalde a priori verdeling van kansen op een aantal verschillende toestanden, kan men nu met behulp van de conditionele kansen op eenvoudige wijze de a posteriori kansen op elk van de mogelijke aandoeningen (de voorspellende waarden,  $P(D_x|T+)$ ) berekenen (43, 44).

Bij ordinale en continu verdeelde variabelen kan het afkappunt (dat een kunstmatige dichotomisering van negatieve en positieve uitslagen markeert) gevarieerd worden. Sensitiviteit en specificiteit zullen dan veranderen. We komen hierop in de volgende paragraaf terug. In zo'n geval kan men het optimale afkappunt trachten vast te stellen. Dit kan geschieden volgens verschillende methoden waarop wij ook terugkomen.

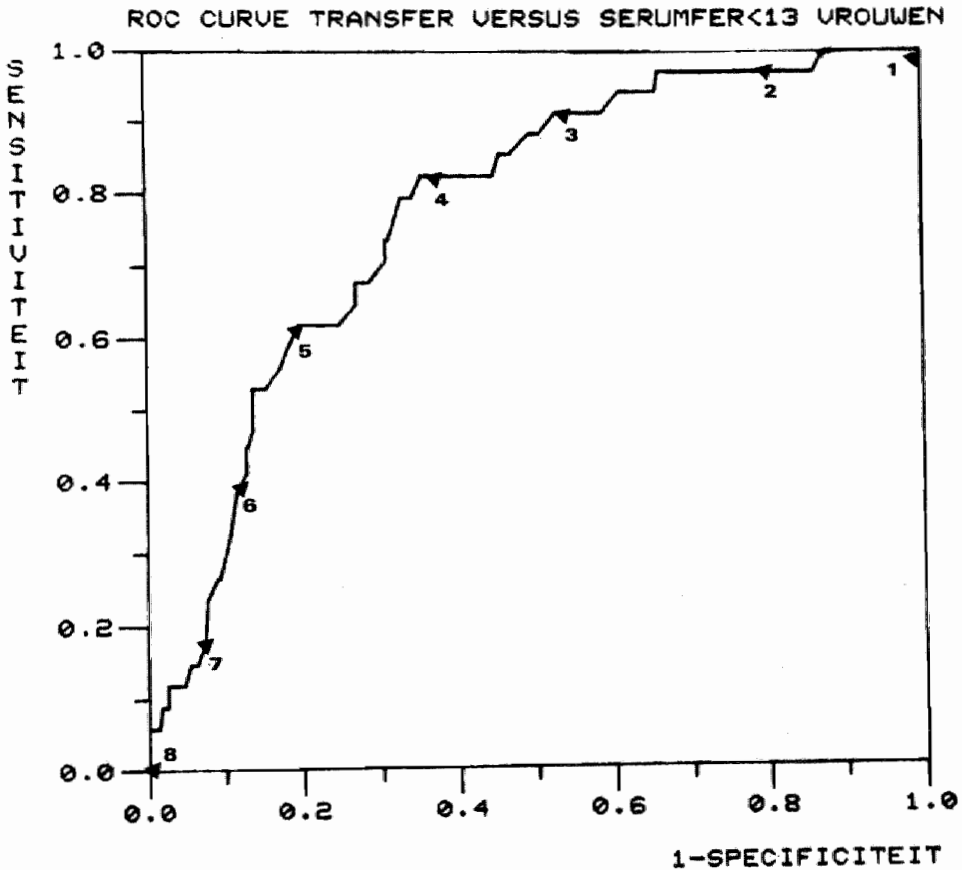
Bij onderzoek ter bepaling van de sensitiviteit en specificiteit of, in het algemeen, van  $P(T+|D_x)$ , is van belang dat er duidelijkheid is over een 'gouden standaard' voor het vaststellen van de bestudeerde ziekten, onafhankelijk van de testuitslag. Wanneer deze niet beschikbaar is of niet toegepast kan worden (wat vaak in de huisartspraktijk het geval is - bijvoorbeeld PA-onderzoek bij alle mensen met buikklachten is ondenkbaar), zal men benaderingen hiervan moeten zoeken. Gedacht kan worden aan (27):

- consensus-beoordeling van de diagnose door een forum van experts die, los van de testuitslag c.q. zonder voorkennis daarover, op grond van het klinisch beeld en zoveel mogelijk op gestandaardiseerde wijze verzamelde aanvullende gegevens tot een eindconclusie komen.
- een follow-up studie, om aan de hand van het klinisch beloop de diagnose definitief te kunnen stellen of uit te sluiten.

#### 4.1.2. ROC-analyse.

Reeds eerder is gezegd, dat voor tests waarbij de testvariabele gemeten wordt op een continue of ordinale schaal geldt, dat het criterium voor het onderscheid tussen positieve en negatieve uitslagen gevari-

afkappunt	transferrine waarde	sensitiviteit	specificiteit
1	2,00	100%	0%
2	2,70	97%	19%
3	3,00	89%	44%
4	3,20	82%	62%
5	3,50	61%	80%
6	3,70	39%	86%
7	3,90	18%	94%
8	6,10	0%	100%



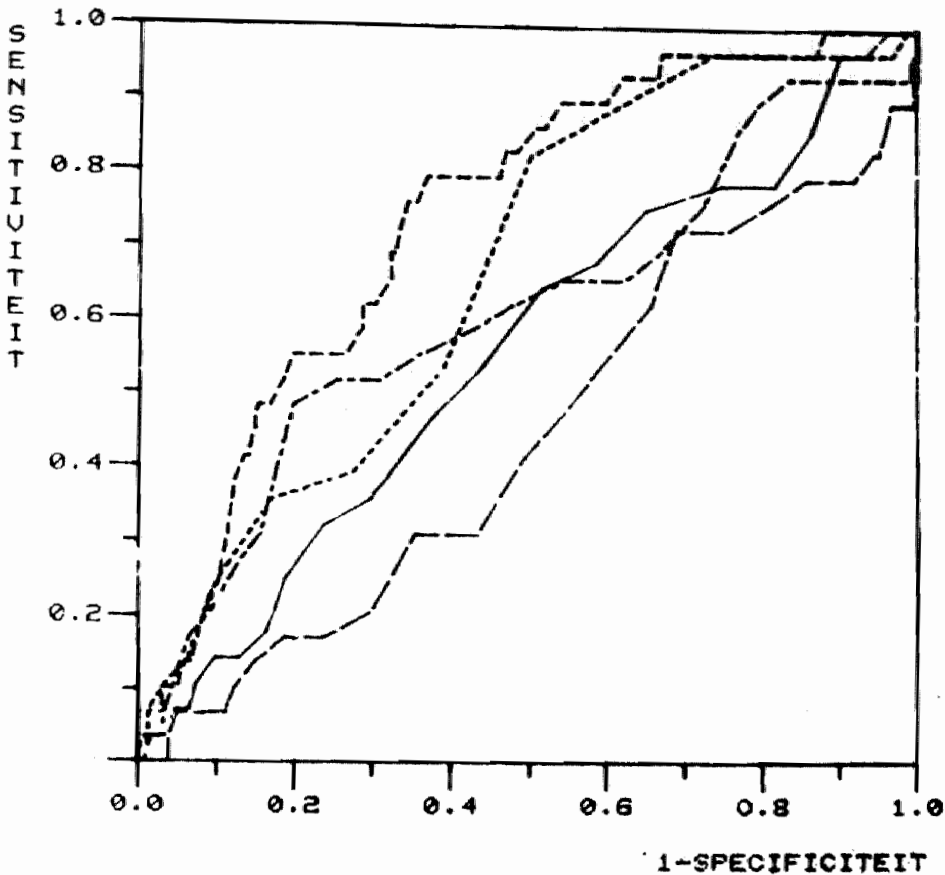
Figuur 1: ROC-curve van de 'test' transferrine concentratie in het veneuze bloed in relatie tot het hebben van een serumferritine concentratie van 12 µg/l of lager.

eerd kan worden in meer of minder stringente richting. Sensitiviteit en specificiteit zullen hierdoor in onderlinge samenhang en in tegengestelde richting veranderen (45). Het is mogelijk om het verloop van sensitiviteit en specificiteit bij een dergelijke variatie van het afkappunt grafisch weer te geven in een zogenaamde Receiver Operating Characteristic (ROC) curve. Daarbij wordt de kans op positieve uitslagen bij zieken (de sensitiviteit) uitgezet tegen de kans op positieve uitslagen bij niet-zieken (100%-specificiteit). In figuur 1 geven wij een voorbeeld op grond van eigen gegevens over 268 patiënten uit de huisartspraktijk. Bij 34 vrouwen met een serumferritine  $\leq 12 \mu\text{g/l}$  en 234 vrouwen met een serumferritine  $> 12 \mu\text{g/l}$  werden de sensitiviteit en specificiteit van het transferrine ten aanzien van 'ijzergebrek' (gedefinieerd door een serumferritinegehalte  $\leq 12 \mu\text{g/l}$ ) bepaald voor een serie afkappunten.

Het is gemakkelijk na te gaan dat naarmate een ROC-curve dichter ligt bij de diagonaal van de grafiek, het onderscheidend vermogen van de test afneemt. Immers als: de sensitiviteit = (100%-de specificiteit) (en dat geldt voor alle punten op de diagonaal), dan is de kans op een positieve uitslag bij zieken even groot als de kans op een (fout-)positieve uitslag bij niet-zieken. Goede tests (met een goed onderscheidend vermogen) kenmerken zich door een snel steil oplopende curve die ver van de diagonaal verloopt.

De vorm van ROC-curves van diverse tests kan dus behulpzaam zijn bij het selecteren van tests met het beste discriminerend vermogen, nog los van de keuze van een bepaald afkappunt.

In het kader van ons onderzoek 'onverklaarde moeheid en Hb-gehalte', was het ons mogelijk de ROC-curves op te stellen voor diverse laboratoriumparameters ten aanzien van ijzergebrek, zoals die gevonden worden in de huisartspraktijk. Voor ijzergebrek werd weer als criterium gedefinieerd een serumferritinegehalte van  $12 \mu\text{g/ml}$  of lager. Voor dit doel werd een computerprogramma geschreven door Drs. S. Roos (Medische Informatica en Statistiek, Rijksuniversiteit Limburg). In figuur 2 zien we het resultaat. De test 'transferrine' heeft het grootste onderscheidend vermogen ten aanzien van de herkenning van vrouwen met een serumferritine concentratie van  $12 \mu\text{g/ml}$  of lager.



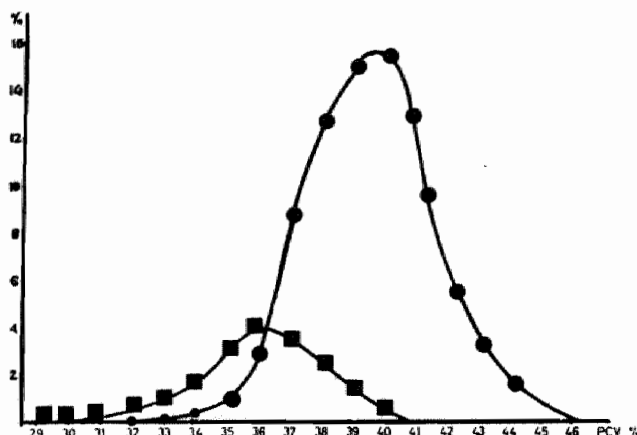
Figuur 2: ROC-curves van diverse laboratorium parameters ten aanzien van de diagnostiek van ijzergebrek (d.w.z. een serumferritine gehalte van 12  $\mu\text{g/l}$  of lager).

- : Hb
- : serumijzer verzadiging
- - - - - : serumijzer
- . - . - : transferrine
- ..... : MCV

Omdat sensitiviteit en specificiteit gevoelig kunnen zijn voor verschillen tussen populaties, kan dit ook gelden voor ROC-curves. Dat wil zeggen: In de huisartspraktijk kunnen ROC-curves er anders uitzien dan in de contactpopulatie van de specialist. Het is mogelijk de in een verwezen populatie vastgestelde ROC-curve te corrigeren voor het verwijsgedrag indien men een schatting heeft van de prior kans op



ziekte in de contactpopulatie van de huisarts, en bovendien wordt aangenomen dat bij verwijzing niet dan alleen via de testuitslag op ziekte geselecteerd wordt (46, 47). Deze aanname is lang niet altijd in overeenstemming met de praktijk. In hoofdstuk 21 komen wij terug op het effect van verwijzingspatronen op het onderscheidend vermogen van diagnostische tests.



Figuur 3: Verdeling van de hematocriet bij vrouwen met (■) en zonder (●) ijzergebrek (steekproef uit de algemene bevolking).

Om informatieverlies te voorkomen kan men ook afzien van het definiëren van één afkappunt, en voor iedere mogelijke waarde van de testparameter de kans op ziekte aangeven (zie figuur 3 en tabel 4 met als voorbeeld de hematocriet in relatie tot ijzergebrek, ontleend aan Garby (9)). Deze kans neemt dan in het algemeen toe als de waarde van de parameter in de 'pathologische' richting verschuift. Ook in de praktijk zal men, vaak reeds intuïtief, een bepaalde betekenis toekennen aan de precieze gemeten waarde: een bloedsuiker van 20 mmol/l betekent iets anders dan een bloedsuiker van 9.0 mmol/l, hoewel beide als 'verhoogd' zijn te beschouwen. Analooq aan wat opgemerkt werd onder paragraaf 2 zou men bij het rapporteren van onderzoekresultaten betreffende de validiteit van tests zich, als het om continue variabelen gaat, niet moeten beperken tot het geven van het percentage waarden

dat onder respectievelijk boven één bepaald afkappunt valt. De complete distributies zouden weergegeven moeten worden voor zieken en niet-zieken. Een diagnost zou dan, voor een bepaalde gemeten waarde zelf, een meer adequate schatting van de kans op ziekte kunnen maken (zie weer het voorbeeld van hematocriet, figuur 3, en paragraaf 4.1.3.).

Tabel 4: Kans op ijzergebrek bij Zweedse vrouwen (steekproef uit de algemene bevolking) bij bepaalde waarden van de hematocriet (zie ook figuur 3).

hematocriet	verhouding "ijzergebrek" : "geen ijzergebrek"	kans op ijzergebrek (%)
< 31	1 : 0	100,0
32	0,51 : 0,08	86,4
33	0,93 : 0,23	80,2
34	1,49 : 0,46	76,4
35	2,79 : 0,92	75,2
36	3,63 : 2,58	58,5
37	3,17 : 8,07	28,2
38	2,33 : 11,67	16,6
39	1,30 : 13,75	8,6
40	0,47 : 14,21	3,2
> 41	0 : 1	0,0

#### 4.1.3. Likelihoodratio (aannemelijkheidsquotiënt) en de oddsratio.

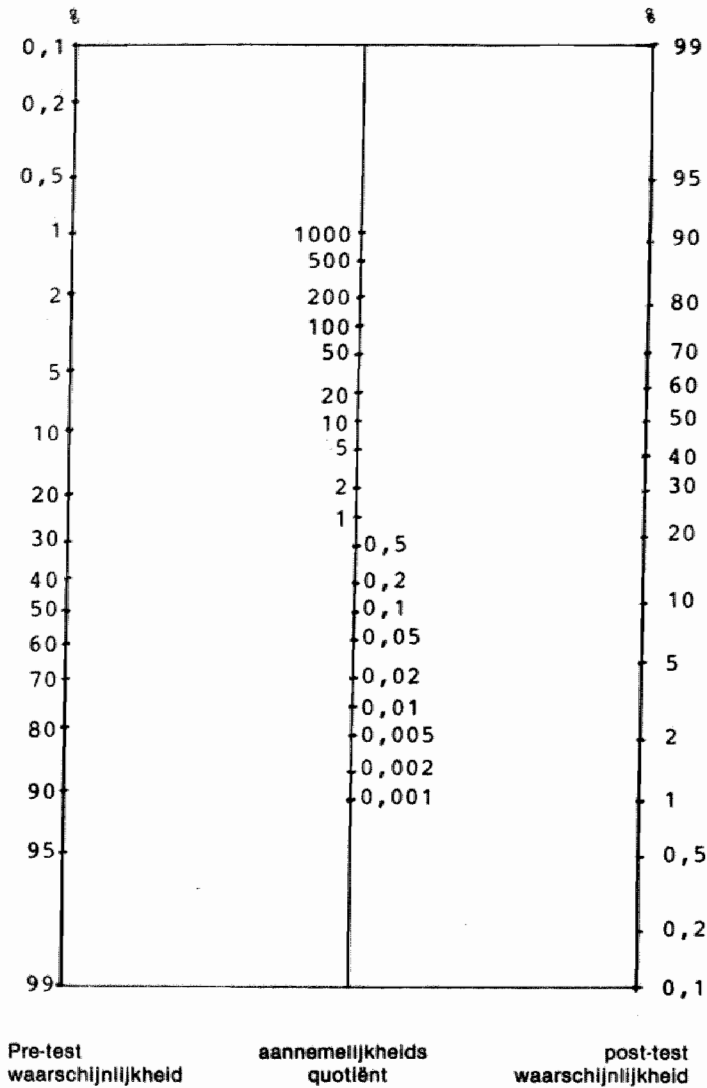
De likelihoodratio is het quotient van de kans dat een bepaald symptoom of een bepaalde testuitslag voorkomt bij zieken, en de overeenkomstige kans bij niet-zieken. Men kan haar voor positieve respectievelijk negatieve uitslagen noteren als (48):

$$LR^+ = \frac{P(T^+|D^+)}{P(T^+|D^-)}, \text{ ofwel: } \frac{\text{sensitiviteit}}{100\% - \text{specificiteit}}$$

en:

$$LR^- = \frac{P(T^-|D^+)}{P(T^-|D^-)}, \text{ ofwel: } \frac{100\% - \text{sensitiviteit}}{\text{specificiteit}}$$

Uit de definitie en de formules is direct af te leiden dat een likelihoodratio van 1.0 betekent dat de test waardeloos is, c.q. geen onderscheidend vermogen heeft. Hoe meer de LR van 1.0 afwijkt (variërend van 0.0 tot oneindig), des te beter is het onderscheidend vermogen.



Figuur 4: Nomogram voor de relatie tussen prevalentie (a priori kans), aannemelijkheidsquotiënt en a posteriori kans. De a posteriori kans is te vinden door met een liniaal vanaf de betreffende "pre-test waarschijnlijkheid", via de relevante waarde voor het aannemelijkheidsquotiënt, de as van de "post-test waarschijnlijkheid" te snijden (49).

In beide maten is de informatie van sensitiviteit en specificiteit samengevat.

De likelihoodratio heeft het voordeel dat zij rechtstreeks gebruikt kan worden in de berekening van de a posteriori kans, door middel van een bepaalde notatie van de formule van Bayes (48).

$$\frac{P(D^+|T^+)}{P(D^-|T^+)} = LR^+ \cdot \frac{P(D^+)}{P(D^-)}, \text{ ofwel: } P(D^+|T^+) = \frac{1}{1 + \frac{1}{LR^+ \times \frac{P(D^+)}{1-P(D^+)}}}$$

en:

$$\frac{P(D^-|T^-)}{P(D^+|T^-)} = \frac{1}{LR^-} \cdot \frac{P(D^-)}{P(D^+)}, \text{ ofwel: } P(D^-|T^-) = \frac{1}{1 + LR^- \times \frac{P(D^+)}{1-P(D^+)}}$$

Op deze wijze kan men dus, gegeven een bepaalde LR, de a posteriori kans berekenen uit de a priori kans. Men kan ook gebruik maken van een nomogram (49) (figuur 4). Voor diverse symptomen en testuitslagen zijn waarden van likelihoodratio's gepubliceerd, dan wel af te leiden uit gepubliceerde waarden van sensitiviteit en specificiteit (49). In tabel 5 geven wij een aantal likelihoodratio's voor de BSE ten aanzien van diverse ziekten (berekend uit tabel 3) en voor transferrine ten aanzien van ijzergebrek (voor diverse afkappunten). Bij de generalisatie van dergelijke gegevens zij men overigens voorzichtig: het is niet zonder meer te verwachten dat likelihoodratio's stabiel zijn over verschillende populaties. Factoren als de mate van ontwikkeling van de ziekte en de mate van selectie ten gevolge van het verwijsgedrag van huisartsen zijn van invloed.

Het is aantrekkelijk om één samenvattende maat voor het onderscheidend vermogen te hebben, uitgaande van sensitiviteit en specificiteit. Op dit punt aangekomen, blijkt dat wij toch nog met twee maten te maken hebben, namelijk  $LR^+$  en  $LR^-$ . Men kan dit oplossen door, als samenvattende maat te kiezen voor het quotient van  $LR^+$  en  $LR^-$ . Dit is de zogenaamde oddsratio (OR), die met name in de etiologische epidemiologie gebruikt wordt. In dit geval kan men haar als volgt definiëren:

$$OR = \frac{\text{Terecht Positieven}}{\text{Fout Negatieven}} \bigg/ \frac{\text{Fout Positieven}}{\text{Terecht Negatieven}}$$

$$= \frac{\text{sensitiviteit}}{(100\% - \text{sensitiviteit})} \times \frac{\text{specificiteit}}{(100\% - \text{specificiteit})} = \frac{LR^+}{LR^-}$$

Tabel 5: Likelihoodratio's voor de BSE ten aanzien van diverse aandoeningen (ontleend aan tabel 3) en voor transferrine ten aanzien van "ijzergebrek" voor diverse afkappunten (ontleend aan figuur 1).

	afkappunt	aandoening	referentie	LR+	LR-	OR
<u>BSE</u>	9	"ziekte", algemeen	33	9,2	0,19	49,4
	30	artritis	35	2,0	0,89	2,3
		osteoartrose		1,7	0,93	1,8
		reumafactor+		2,7	0,79	3,5
	"low"	afwijkingen plasma-eiwit	36	1,7	0,11	15,5
	"intermediate"			3,7	0,22	17,3
	"high"			7,6	0,35	21,5
	14	adnexitis	37	0,6	1,09	0,58
	10	kanker	39	1,0	1,30	0,7
	20			1,0	1,00	1,0
	30			1,1	0,89	1,2
<u>Transferrine</u>	2,0	ijzergebrek	fig. 3	1,0	*	*
	2,7	(serumferri-		1,2	0,16	7,5
	3,0	tine 12 µg/l		1,6	0,25	6,4
	3,2			2,2	0,29	7,4
	3,5			3,1	0,49	6,3
	3,7			2,8	0,71	3,9
	3,9			3,0	0,87	3,4
	4,4			6,9	0,98	7,1
	6,1			*	1,0	*

\* niet te berekenen

De OR is gelijk aan 1.0 als het diagnosticum geen onderscheidend vermogen heeft en wijkt hiervan sterker af (variërend van 0 tot oneindig) naarmate het onderscheidend vermogen beter is. In de laatste kolom van tabel 5 is de waarde van de OR ter oriëntatie toegevoegd. In het algemeen blijkt dat de OR zeer hoog wordt als men nadert tot de grenzen van het overgangsgebied waarbinnen de test-uitslagen van zieken en niet-zieken elkaar overlappen: de noemer van de breuk nadert dan tot 0. Onder de ondergrens en boven de bovengrens van dit overlappingsgebied hebben de OR en respectievelijk de  $LR^-$  en  $LR^+$  geen betekenis meer, en hebben ook hun functie verloren: er is immers geen overlap meer tussen zieken en niet-zieken en men mag dus zeker van zijn zaak zijn.

Over het nut van de odds-ratio als samenvattende maat voor het onderscheidend vermogen kan men van mening verschillen. Het voordeel is de 'samenvatting'. Het nadeel is, dat men haar niet zonder meer kan gebruiken voor het berekenen van a posteriori kansen. Immers, dezelfde oddsratio kan gelden voor verschillende waarden van sensitiviteit en specificiteit respectievelijk  $LR^+$  en  $LR^-$ . Zij is ook een functie van de testvariabele zoals in het tweede deel van tabel 5 te zien is. Voor dichotome test-variabelen is de interpretatie van de odds-ratio meer eenduidig dan voor continue variabelen, daar er geen afhankelijkheid is van het (immers niet te variëren) afkappunt. In dat geval is zij dan ook meer geschikt als samenvattende maat voor het onderscheidend vermogen.

Voor continue testvariabelen is er nog een andere benadering, die tevens het meest efficiënte gebruik van de diagnostische informatie inhoudt, namelijk: het verstrekken van de likelihoodratio die behoort bij de gemeten (klasse van) waarde(n) van de testuitslag, zonder hierbij dus te hoeven dichotomiseren in negatieve en positieve uitslagen. De likelihoodratio is dan:

$$LR_x = \frac{P(T_x | D^+)}{P(T_x | D^-)} = \frac{\text{kans op testuitslag } x \text{ bij zieken}}{\text{kans op testuitslag } x \text{ bij niet-zieken}}$$

Op grond van de a priori kans en  $LR_x$  kan men dan, zoals eerder beschreven, de a posteriori kans berekenen maar nu meer precies, uitgaande van de 'exacte' testuitslag x.

Tabel 6: Likelihoodratio's van transferrine ten aanzien van ijzergebrek (serumferritine  $\leq 12 \mu\text{g/l}$ ), berekend voor verschillende klassen van waarden.

transferrine waarde	ijzergebrek	geen ijzergebrek	$LR_x$
-2.49	0	20	0,0
2.50-3.24	6	124	0,33
3.25-3.99	24	78	2,11
4.00-4.74	2	11	1,25
4.75-5.49	1	1	13,8*
5.50 +	1	0	
totaal	34	234	

\* i.v.m. de kleine aantallen werden deze categorieën tezamen genomen.

In tabel 6 hebben wij, ter illustratie,  $LR_x$  van diverse klassen van transferrinewaarden gegeven voor het diagnostiseren van 'ijzergebrek' (c.q. een serumferritine waarde onder de norm). Omdat echter ook een 'klasse' nog geen precies gemeten waarde is, is het wenselijk een statistisch model op te stellen, op grond van empirische data, voor  $LR_x$  als functie van de testvariabele, waaruit  $LR_x$  voor iedere waarde volgt. Voor transferrine, dat bij benadering normaal verdeeld is, konden wij dit doen door middel van logistische regressie, waarbij wij het intercept gelijk maakten aan de natuurlijke antilogaritme van de prior odds. Om dit te bereiken dient men de onafhankelijke variabele  $x$  te transformeren tot  $x-x_0$ , waarbij  $x_0$  dié waarde van  $x$  is waarbij de posterior disease-odds gelijk is aan de prior disease-odds. Met andere woorden:  $x_0$  is juist die testuitslag die de kans op ziekte niet groter en niet kleiner maakt. Dan geldt immers (zie ook appendix bij hoofdstuk 8):

$$\text{posterior odds} = e^{a+\beta(x-x_0)} = e^a \times e^{\beta(x-x_0)}$$

$$\text{en posterior kans} = \frac{1}{1 + \frac{1}{e^{a+\beta(x-x_0)}}}$$

waarbij:  $x_0$  = waarde van  $x$  waarbij geldt dat posterior odds = prior odds.

$$e^a = \text{prior disease-odds.}$$

$$e^{\beta(x-x_0)} = LR_x$$

Ter illustratie geven wij in tabel 7 het model voor transferrine als test voor ijzergebrek, met de eruit volgende waarden voor  $LR_x$ , de posterior odds, en de posterior kans.

Op grond van het beloop van genoemde functie van  $LR_x$  over het overlappingstraject kan men het onderscheidend vermogen van de test beoordelen zonder de testuitslag te hoeven dichotomiseren (wat bij een ROC-curve wel gebeurt).

Tabel 7: Logistisch model voor de berekening van  $LR_x$ , de likelihood-ratio van bepaalde transferrine waarden<sup>x</sup> ten aanzien van ijzergebrek, en de hieruit volgende posterior kansen.

model:	coëfficiënt	standard error
intercept	- 1.94	0.206
(transferrine - 3.36)	1.89	0.415

dus: posterior odds =  $e^{-1.94 + 1.89 (x-3.36)}$

waarbij :  $e^{-1.94}$  = prior odds  
 $e^{1.89(x-3.36)}$  =  $LR_x$   
 3.36 = de transferrine waarde waarvoor geldt, dat de posterior odds = prior odds

Toepassing:

transferrine waarde	$LR_x$	posterior odds	posterior kans
2.50	0.20	0.03	0.03
3.00	0.51	0.07	0.07
3.36	1.00	0.14	0.13
3.50	1.30	0.19	0.16
4.00	3.35	0.48	0.33
4.50	5.38	0.77	0.44
5.00	22.19	3.12	0.76
5.50	91.57	13.16	0.93

Er is in de praktijk veel voor te zeggen om bij het zoeken naar normen niet één afkappunt te kiezen, maar twee. Onder het laagste afkappunt (waar de sensitiviteit = 100%) is de voorspellende waarde ten aanzien van niet-ziekte 100%, en boven het hoogste (waar de specificiteit = 100%) is de voorspellende waarde ten aanzien van ziekte 100%. Tussen beide afkappunten zijn we onzeker, en zijn de likelihoodratio, de prior kans en de voorspellende waarde van belang bij het reduceren van de onzekerheid (27).



4.1.4. Het bepalen van het afkappunt waarbij de te verwachten toegenomen zekerheid van een ziekte als gevolg van de test maximaal is.

De predictieve waarden (a posteriori kansen)  $P(D^+|T^+)$  en  $P(D^-|T^-)$  van een testuitslag op zichzelf zijn onvoldoende om de informativiteit van een test te beoordelen. Het gaat om het verschil tussen:

- de a priori kans en de a posteriori kans van een positieve test:  $P(D^+|T^+)$  respectievelijk  $P(D^+)$ , en tussen:
- de a priori kans en de a posteriori kans van een negatieve test:  $P(D^-|T^-)$  respectievelijk  $P(D^-)$ .

Men kan ook spreken over de mate van toename in zekerheid dankzij het verrichten van de test. Deze toename kan in diverse 'maten' uitgedrukt worden:

a) netto 'toename':  $P(D^+|T^+) - P(D^+)$  resp:  $P(D^-|T^-) - P(D^-)$

b) procentuele toename:  
(als percentage van de a priori kans)

$$\frac{P(D^+|T^+) - P(D^+)}{P(D^+)} \quad \text{resp:} \quad \frac{P(D^-|T^-) - P(D^-)}{P(D^-)}$$

c) percentage van maximaal haalbare toename  
(d.w.z. als percentage van het verschil tussen 100% zekerheid en de a priori kans)

$$\frac{P(D^+|T^+) - P(D^+)}{1 - P(D^+)} \quad \text{resp:} \quad \frac{P(D^-|T^-) - P(D^-)}{1 - P(D^-)}$$

d) relatieve toename  
t.o.v. de a priori kans

$$\frac{P(D^+|T^+)}{P(D^+)} \quad \text{resp:} \quad \frac{P(D^-|T^-)}{P(D^-)}$$

e) de verhouding tussen:  
de kans op ziekte  
en niet-ziekte ná:  
resp. vóór toepassing  
van de test.

$$\frac{P(D^+|T^+)}{P(D^-|T^+)} \bigg/ \frac{P(D^+)}{P(D^-)} \quad \text{en:} \quad \frac{P(D^+|T^-)}{P(D^-|T^-)} \bigg/ \frac{P(D^+)}{P(D^-)}$$

De laatstgenoemde maat is onafhankelijk van de a priori kans: zij is namelijk gelijk aan de likelihoodratio van een positieve respectievelijk negatieve test, die we al bespraken.

De eerst vier maten zijn wél afhankelijk van de a priori kans.

Conell en Koepsell behandelen de drie eerste genoemde en spreken in dit verband van: 'measures of gain in certainty from a diagnostic test' (50).

Een positieve test en een negatieve test hoeven niet even frequent voor te komen. Het kan zijn, dat in het geval van een positieve uitslag de kans op ziekte sterk toeneemt ten opzichte van de vooraf kans, maar dat de kans op een positieve uitslag in de te testen populatie zeer laag is. Deze laatste kans is mede bepalend voor de te verwachten toename in zekerheid door het uitvoeren van de test. Indien de kans op een positieve respectievelijk negatieve test vóór het uitvoeren van de test te voorspellen is, kan men de toename van de zekerheid door de test vooraf wegeven voor de kans op een bepaalde uitslag. Aldus kunnen de uitdrukkingen  $a\ t/m\ c$  worden getransformeerd tot 'measures of expected gain in certainty' ten aanzien van respectievelijk een positieve uitslag, een negatieve uitslag en de gecombineerde informatie van positieve en negatieve uitslagen. Voor een gedetailleerde uitwerking verwijzen we naar Conell en Koepsell (50). In het kader van dit hoofdstuk, waarin het gaat om 'normen' voor testuitslagen, zijn vooral de volgende aspecten van belang.

- 1) De verwachte absolute toename in zekerheid (certainty) als gevolg van de te verwachten gecombineerde informatie van positieve en negatieve uitslagen is gelijk aan:

$$E(\Delta c) = 2 \cdot P(D^+) \cdot P(D^-) \cdot \left\{ P(T^+|D^+) + P(T^-|D^-) - 1 \right\}$$

Dit is een functie van  $P(D^+)$ , de a priori kans op ziekte.

Deze is maximaal als:  $P(D^+) = 0.5$ , dat wil zeggen als de a priori kans 50% is: er is sprake van maximale twijfel, er is een 'fifty-fifty' kans op ziekte. (Zie ook McNeill et al. (51)).

en minimaal als:  $P(D^+)$  nadert tot 0 of tot 1.

Als functie van sensitiviteit  $P(T^+|D^+)$  en specificiteit  $P(T^-|D^-)$  wordt  $E(\Delta c)$  maximaal als  $P(T^+|D^+) + P(T^-|D^-) - 1$  maximaal is. Dat betekent dus dat, bij een continu verdeelde testvariabele het afkappunt zo gekozen zou moeten worden dat de som van sensitiviteit en specificiteit maximaal is.

- 2) De maten voor de verwachte toename in zekerheid, als percentage van de reeds bestaande zekerheid vóór de test ( $E(\Delta c)$ ) respectievelijk als percentage van de maximaal mogelijke toename ( $E(\Delta c_{\max})$ ), voor de gecombineerde informatie van positieve en negatieve uitslagen zijn beide gelijk aan:

$$E(\Delta c) = E(\Delta c_{\max}) = P(T^+|D^+) + P(T^-|D^-) - 1$$

Deze maten zijn nu dus onafhankelijk van de prevalentie, en weer maximaal als de som van de sensitiviteit en de specificiteit maximaal is.

De conclusie is dat, als men streeft naar maximale winst in zekerheid ten gevolge van het uitvoeren van een test, men het afkappunt moet kiezen waarvoor geldt dat de som van de sensitiviteit en de specificiteit maximaal is. Op de ROC-curve betreft dit het punt waarop de afstand tot de diagonaal maximaal is (50). Indien de uitslagen bij zieken en niet-zieken een 'normale' verdeling met gelijke spreiding volgen, is dit het punt waarbij sensitiviteit en specificiteit gelijk aan elkaar zijn.

Het opmerkelijke van deze benadering is, dat duidelijk wordt dat een maximale likelihoodratio van een positieve test niet betekent dat de toegenomen zekerheid (op grond van de gecombineerde informatie van positieve en negatieve uitslagen) ten gevolge van de test ook maximaal is. De maximale likelihoodratio vindt men immers als de specificiteit nadert tot 100%: de likelihoodratio nadert dan tot oneindig, ook als de sensitiviteit zeer klein is.

Ook hier geldt overigens weer dat de gebruikte maten afhankelijk zijn van kenmerken van de bestudeerde populatie, met name van de mate van

selectie ten aanzien van de testuitslag en van de mate van ontwikkeling van de ziekte.

#### 4.2. Test en lange termijnprognose.

In het voorgaande is gesproken over het verband tussen testuitslagen en de aanwezigheid van ziekte op het moment van het uitvoeren van de test. Het is echter ook mogelijk om de testuitslag te gebruiken als voorspeller van toekomstige ontwikkelingen. Bekende voorbeelden zijn de bloeddruk en het cholesterolgehalte. Het hebben van een hoge bloeddruk betekent uiteraard niet dat men een cerebrovasculair accident heeft. Wel kan men er vanuit gaan dat een hoge bloeddruk de kans op het krijgen van een cerebrovasculair accident in de toekomst verhoogt (52, 53). Hetzelfde geldt voor het verband tussen het cholesterolgehalte en het myocardinfarct (52, 54, 55). Men ziet dan een hoge bloeddruk en een hoog cholesterol niet zozeer als diagnostica, maar als risico-indicatoren ten aanzien van bepaalde mogelijke verwickelingen in de toekomst.

Hierbij zijn de volgende kanttekeningen te maken:

- Een risico-indicator kan al dan niet een causale betekenis hebben. Het hebben van gele vingers heeft, in tegenstelling tot roken, geen causale betekenis ten aanzien van longkanker. Toch kan het, evenals roken, gelden als een risico-indicator. Tests kunnen een prognostische betekenis hebben, ook al komen ze in de oorzaak-gevolg keten van de ziekte niet voor.
- Het vinden van een risico-indicator kan, evenals een diagnosticum, aanleiding geven tot interventie. Daarbij kan de risico-indicator zélf het doelwit zijn van de beïnvloeding omdat hij geacht wordt een causale rol te spelen én bovendien beïnvloedbaar is, (zoals bijvoorbeeld hoge bloeddruk). Vaak echter is de 'test' zelf niet te beïnvloeden (zoals de resultaten van genetisch onderzoek, en de rhesus-factor) maar maakt het resultaat ons alert, teneinde eventueel te kunnen anticiperen (amnionpunctie; controle op rhesusantistoffen). Soms is wel een verhoogd risico vast te stellen, maar kan zal men (vooralsnog) nóch interventie kunnen plegen, nóch hierop kunnen anticiperen. Een voorbeeld zou kunnen zijn het positieve verband tussen een matig hoge hematocriet en de kans op een hartinfarct (56, 57).

Men zou kunnen spreken van het diagnostiseren van risico's met behulp van risico-indicatoren, en het behandelen van risico-indicatoren voor zover ze causaal van betekenis zijn.

Een maat om de karakteristieken van de test in prognostische zin te beschrijven is: het relatief risico. Dit komt in het verband van deze bespreking overeen met:

de kans op het krijgen van de aandoening bij 'test-positieven',  
in een bepaalde periode  
de kans op het krijgen van de aandoening bij 'test-negatieven',  
in dezelfde periode.

Zelf verrichtten wij een onderzoek naar de betekenis van het hemoglobinegehalte, c.q. de hematocriet voor het risico op het krijgen van een hartinfarct. Wij maakten hierbij gebruik van een patiënt-controle methode, waarbij het Hb vóór het optreden van het infarct werd gemeten. Wij konden de bevindingen van diverse onderzoekers (56, 57), dat er sprake is van een verhoogd risico bij een verhoogde hematocriet, bevestigen. Wij vonden, dat in de leeftijdsperiode van 50-60 jaar, mannen die een hematocriet hebben van 46 of hoger, over het geheel genomen een 2.7 maal zo grote kans hebben op het krijgen van een infarct als mannen met een hematocriet lager dan 46. Er blijft hierover overigens nog een controverse, omdat anderen het verband tussen een matig hoog hemoglobinegehalte of hematocriet en hartinfarct niet of nauwelijks vonden (58-61). In hoofdstuk 23 doen wij verslag van dit onderzoek.

Het relatieve risico heeft als pendant in de diagnostiek: het quotiënt van de voorspellende waarde van een kenmerk of positieve testuitslag en het complement van de voorspellende waarde van de afwezigheid van het kenmerk of de negatieve uitslag. Het ligt voor de hand om na te gaan of er andersom ook diagnostische parameters voor de prognostiek bruikbaar zijn, uitgaande van bepaalde risico indicatoren. Hierbij komt in aanmerking de likelihoodratio (LR), hier te benoemen als de prognostische likelihoodratio van een prognostisch kenmerk of een positieve testuitslag. Men kan deze bepalen op grond van cohortstudies met een bepaalde follow-up duur  $\Delta t$  of op grond van patiënt-controle

onderzoeken (cumulatieve incidentie-methode). Zij is te omschrijven als: de kans dat personen, die de ziekte na een periode  $\Delta t$  gekregen hebben, op  $t_0$  (bij het begin van de follow-up) een positieve uitslag hadden, gedeeld door de overeenkomstige kans bij personen die na de periode  $\Delta t$  de ziekte niet gekregen hebben. Met behulp van deze ratio kan men de kans berekenen dat personen met een positieve uitslag na een periode  $\Delta t$  de ziekte gekregen hebben, indien men tevens beschikt over een schatting van de "a priori kans" dat de ziekte optreedt (de geschatte kans zonder voorkennis over de test-uitslag). Deze a priori kans is gelijk aan de cumulatieve incidentie over de periode  $\Delta t$  voor de gehele bestudeerde populatie, dus met de personen met én zonder het kenmerk of de positieve uitslag.

We krijgen dan de volgende vergelijkingen:

$$\frac{(CI_{\Delta t}|F+)}{(1-CI_{\Delta t}|F+)} = \frac{P_t(F+|D+)}{P_t(F+|D-)} \cdot \frac{CI_{\Delta t}}{1-CI_{\Delta t}}$$

waarbij:

$$\frac{CI_{\Delta t}|F+}{1-(CI_{\Delta t}|F+)} =$$

$\frac{\text{cumulatieve incidentie na periode } \Delta t \text{ bij patiënten met kenmerk } F}{1 - \text{cumulatieve incidentie na periode } \Delta t \text{ bij patiënten met kenmerk } F}$

$\frac{P_t(F+|D+)}{P_t(F+|D-)} =$  likelihoodratio, bepaald uit een cohort-onderzoek of patiënt-controle onderzoek na afloop van  $\Delta t$ .

$\frac{CI_{\Delta t}}{1-CI_{\Delta t}} = \frac{\text{cumulatieve incidentie voor de gehele patiëntengroep na afloop van } \Delta t}{1 - \text{cumulatieve incidentie voor de gehele patiëntengroep na afloop van } \Delta t}$  (= 'prior disease-odds')

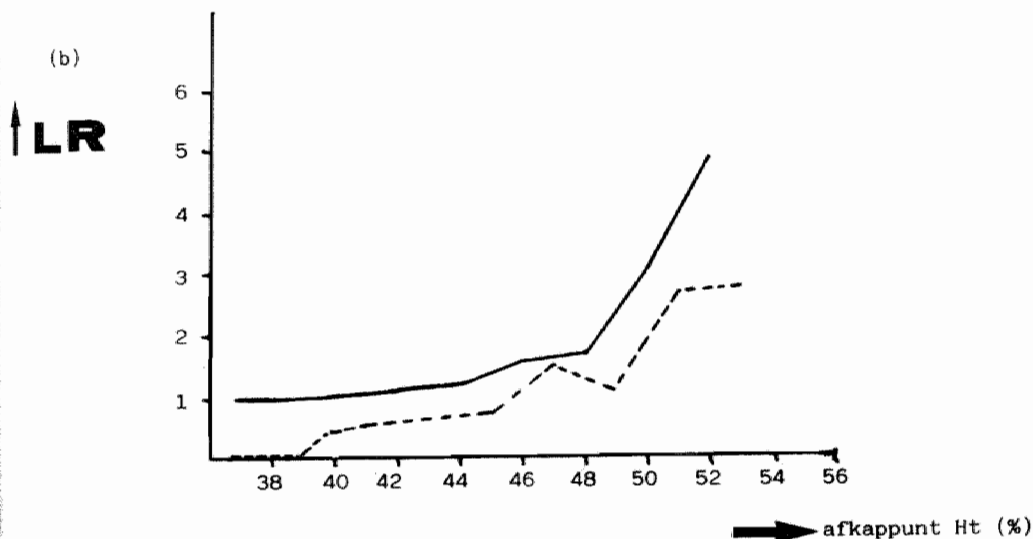
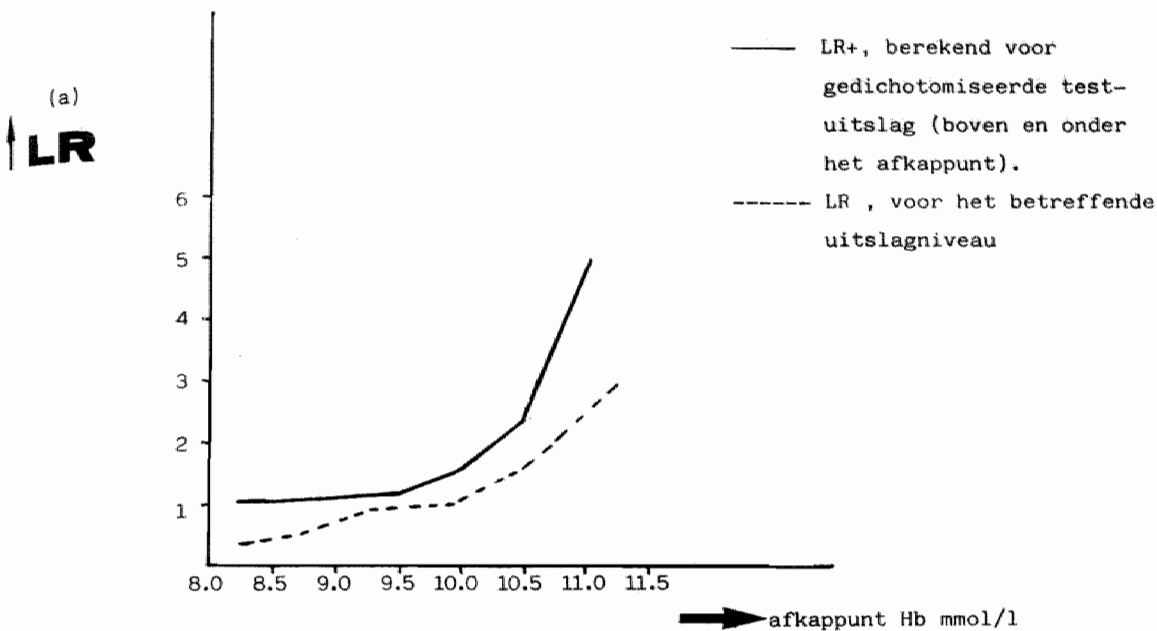
De vergelijking is equivalent aan:

posterior odds = likelihoodratio x prior odds.

Evenals in de diagnostische situatie, moet de eis worden gesteld dat de 'prognostische likelihoodratio' bepaald is bij zieken en niet-zieken uit een bronpopulatie die op relevante kenmerken representatief is voor de populatie waarop men de prognostische schatting wil toepassen.

De patiënten-groep en de niet-zieken moeten ook representatief zijn voor de personen die na  $\Delta t$  de ziekte wel, respectievelijk niet hadden. Een praktisch probleem dat hierbij kan optreden - en wel des te meer naarmate  $\Delta t$  langer is - is selectieve uitval van patiënten en wellicht vooral die patiënten die kenmerk F bezitten. De prognostische likelihoodratio zal hierdoor onderschat kunnen worden. Dit zou ook gebeurd kunnen zijn ten aanzien van ons patiënt-controle onderzoek naar het verband tussen het optreden van een eerste hartinfarct en de hematocriet respectievelijk het hemoglobinegehalte, waarover wij elders in deze bundel verslag uitbrengen. Met deze kanttekening presenteren wij ter illustratie van deze paragraaf het verloop van de prognostische likelihoodratio's van de hematocriet en het hemoglobinegehalte voor mannen van 50-60 jaar, voor diverse afkappunten en precieze waarden, zoals berekend op grond van onze resultaten (ongematchte analyse) (figuur 5). Men ziet dat deze likelihoodratio's vooral gaan stijgen als zij hoger worden dan een hematocriet van 48 en een hemoglobine van 10.0 mmol/l.

Het voordeel van de prognostische likelihoodratio ten opzichte van het relatief risico en de odds-ratio is, dat indien deze likelihoodratio eenmaal bekend is (eventueel voor verschillende follow-up duren) men slechts de geschatte a priori kans op het ontwikkelen van de ziekte hoeft aan te passen op grond van het nieuwe diagnostische gegeven. Men hoeft niet meer uit te gaan van de incidentie bij mensen zonder het kenmerk. Dit is vooral een voordeel bij het formuleren van de prognose van individuele patiënten, voor wie vaak uitgegaan wordt van een meer op de persoon afgestemde a priori kans op bepaalde verwikkelingen in de toekomst. Bovendien kan men, vooral bij continue variabelen lang niet altijd spreken van personen zonder het kenmerk: iederéén heeft een bloeddruk, een cholesterol- of een hemoglobine waarde. Het is dan beter om als referentiepunt te kiezen de overall (a priori) kans op ziekte, nog niet rekening houdend met het kenmerk of testuitslag. Verder is een voordeel, dat de LR en de a priori kans op het krijgen van de ziekte in principe van elkaar onafhankelijk zijn, terwijl het relatief risico mede beïnvloed wordt door de a priori kans: als deze zeer hoog is, dan nadert het relatief risico tot 1.0. Uiteindelijk



Figuur 5 Het verband tussen de prognostische likelihoodratio en het Hemoglobinegehalte (Hb) respectievelijk de Hematocriet (Ht) bepaald op grond van een patiënt-controle onderzoek naar het verband tussen het optreden van een eerste hartinfarct en Hb respectievelijk Ht, bij mannen van 50-60 jaar.



moet op grond van empirische gegevens worden vastgesteld welke maat beter c.q. het meest constant is bij toepassing in verschillende populaties en in relatie tot de follow-up duur. Dit kan voor verschillende kenmerken en aandoeningen anders blijken te liggen.

Het verband tussen de schatting van de kans op het krijgen van de ziekte door middel van de Bayesiaanse methode en de schatting van deze kans met behulp van een logistisch model is eenvoudig. Immers, het gaat in beide gevallen om dezelfde einduitkomst, namelijk:  $CI_{\Delta t} | F^+$ . Deze is gelijk aan:

$$\frac{1}{1 + \frac{1}{e^{a + \beta x}}} = \frac{1}{1 + \frac{1}{LR_{F^+} \times \text{prior disease-odds}}} (= CI_{\Delta t} | F^+)$$

(volgens het logistische model)                      (volgens het Bayesiaanse model)

Zodat:

$$a + \beta x = \ln (LR_{F^+} \times \text{prior odds}).$$

Indien men nu een dichotome onafhankelijke variabele  $x$  transformeert tot  $(x - x_0)$ , waarbij  $x_0$  die (virtuele) waarde van  $x$  is (tussen 0 en 1) waarvoor de prior kans gelijk is aan de posterior kans, dan geldt:

$$CI_t | F^+ = e^{a + \beta (x - x_0)} = e^a \cdot e^{\beta (x - x_0)}$$

$$\begin{aligned} \text{Hierbij is: } e^a &= \text{de prior disease odds } (x=x_0) \\ e^{\beta (x - x_0)} &= \text{de prognostische likelihoodratio} \end{aligned}$$

$$e^{a - \beta x_0} = \text{de posterior disease-odds bij een negatieve test-uitslag of bij afwezigheid van kenmerk } (x=0)$$

$$e^{a + \beta x} = \text{de posterior disease-odds bij een positieve uitslag of bij aanwezigheid van het kenmerk } (x=1)$$

Bij continue x-variabelen is het model op dezelfde wijze toepasbaar. Daarbij kiest men voor  $x_0$  weer die waarde van x waarbij de posterior kans gelijk is aan de prior kans (zie ook de appendix bij hoofdstuk 8 en paragraaf 4.1.3. van dit hoofdstuk).

## 5. Normen in relatie tot het gewicht van de diverse uitkomsten.

Het uiteindelijke doel van het verrichten van diagnostiek is het bijdragen aan de verbetering van de gezondheidstoestand van de onderzochte. Dit impliceert, dat de beslissing om diagnostiek te doen, en de wijze van interpretatie van de uitslagen in verband moeten worden gebracht met de te verwachten invloed op de gezondheidstoestand. Er zijn diverse manieren om dit gezichtspunt uit te werken.

### 5.1. Diagnostiek en behandelbaarheid.

Diagnostiek wordt gelegitimeerd door haar mogelijk gunstige invloed op de prognose, via een naar aanleiding van de diagnostische resultaten in of bij te stellen behandeling. Men kan hieraan een criterium ontleenen voor de interpretatie van testuitslagen:

De uitslag van het diagnostisch onderzoek is positief als deze wijst op een behandelbare aandoening of toestand.

De uitslag is negatief uitgevallen wanneer deze wijst op géén aandoening of op een aandoening die niet - of niet meer - behandelbaar is.

Een voorbeeld is het carotisangiogram. Het kan enerzijds wijzen op operabele stenosen (positief) en anderzijds op afwezigheid van stenosen of op inoperabele stenosen (negatief). Het gaat hierbij om de klinische relevantie van de uitslag voor het beleid.

Hierbij is nog onderscheid te maken tussen 'technisch' behandelbaar en vatbaar voor zinvolle interventie. Immers: een carotisangiogram kan wijzen op een op zich operabele stenose, terwijl het nut van endarterectomie voor de prognose nog niet is aangetoond (zie hoofdstuk 18).

Het is duidelijk dat een op deze wijze gedefinieerd positiviteitscriterium niet gebaseerd is op de sensitiviteit en specificiteit ten aanzien van de aanwezigheid van ziekte, maar op de sensitiviteit en specificiteit (of voorspellende waarde) ten aanzien van de behandelbaarheid van de ziekte. Dat introductie van het begrip 'behandelbaar'

tot andere interpretaties kan leiden, blijkt uit het al eerder aangehaalde onderzoek van Garby (9). Door het geven van een proefbehandeling met oraal ijzer kon hij twee subpopulaties onderscheiden in een steekproef van vrouwen uit de algemene bevolking van Zweden: een (grote) subpopulatie die niet reageerde op ijzertherapie met een stijging van hematocriet en hemoglobine, en een (kleine) subpopulatie die wel reageerde. Hoewel de hematocriet- en hemoglobinewaarden van de wel reagerende groep in het algemeen lager waren dan van de niet-reagerende groep was er een sterke overlap: in beide groepen kwamen veel waarden voor ter weerszijden van het gebruikelijke afkappunt (zie figuur 3).

Stel dat men reeds op geleide van de gemeten hematocriet- of hemoglobinewaarde (dus zonder ijzer bepalingen te verrichten) zou willen besluiten om al dan niet tot ijzerbehandeling over te gaan (hetgeen in de huisartspraktijk niet ongebruikelijk is). De vraag, waar men in zo'n geval het afkappunt zou willen aangeven, reikt verder dan het bereiken van een maximaal onderscheidend vermogen. De beantwoording hangt mede af van het gewicht dat men hecht aan:

- het terecht behandelen (terecht-positief)
- het terecht niet behandelen (terecht-negatief)
- het ten onrechte behandelen (fout-positief)
- het ten onrechte niet-behandelen (fout-negatief).

In 5.2. zullen we dit verder uitwerken.

#### 5.2. De bepaling van het optimale afkappunt van een test met een continu (of ordinaal) verdeelde testvariabele, aan de hand van 'utiliteiten'.

Men kan de consequenties van terecht positieve, terecht negatieve, fout-positieve en fout-negatieve uitkomsten uitdrukken in waarden (utiliteiten), respectievelijk te noemen:

$U_{TP}$ ,  $U_{TN}$ ,  $U_{FP}$  en  $U_{FN}$ .

Met behulp van een eenvoudige beslisboom, de formule van Bayes en de gelijkstelling van de gezamenlijke waarde van alle positieve (TP en FP) aan de gezamenlijke waarde van alle negatieve uitkomsten (FN en TN), is de volgende formule op eenvoudige wijze af te leiden (48).

$$LR_x = \frac{P(T > x | D^+)}{P(T > x | D^-)} = \frac{P(D^-)}{P(D^+)} \times \frac{U_{TN} - U_{FP}}{U_{TP} - U_{FN}}$$

waarbij  $LR_x$  staat voor de likelihoodratio voor testuitslagen  $> x$ , met het optimale afkappunt  $x$  (waarbij de gezamenlijke consequenties van positieve uitslagen even zwaar wegen als die van negatieve uitslagen)

en:  $\frac{P(D^-)}{P(D^+)}$  de inverse is van de prior odds (48).

Uit de formule blijkt dat de waarde  $x$  van het optimale afkappunt (c.q. de optimale likelihoodratio) afhangt van de gekozen utiliteiten van iedere mogelijke testuitkomst. Deze kunnen per arts en/of per patiënt of subgroep van patiënten verschillen.

Ook is te zien dat de waarde  $x$  van het optimale afkappunt afhangt van de a priori kans: hoe groter de a priori kans, des te kleiner  $LR_x$  (d.w.z. des te minder stringent dient het afkappunt te zijn).

Het begrip 'norm'- of grenswaarde verliest nu in feite haar algemeen geldende karakter. Het zal van de specifieke situatie en met name de voorkeuren van de patiënt(en) en de a priori kansen kunnen afhangen welke norm men kiest.

Men kan ook uitgaan van een bepaalde precieze waarde van de testvariabele  $T$ . Zonder nu een (geforceerde en inefficiënte) indeling te maken in positieve en negatieve uitslagen  $T^+$  en  $T^-$  kan men spreken van  $P(D^+|T)$  en  $P(D^-|T)$ : de kans op ziekte, respectievelijk niet-ziekte gegeven de gemeten testuitslag  $T$  (zie ook figuur 3 en 4.1.3.).

Voor welke diagnostische conclusie men kiest (wel of niet ziekte constateren, dan wel tot nader onderzoek overgaan) hangt weer mede af van het gewicht van die conclusie voor de diverse classificatie mogelijkheden van  $T$  ( $U_{TP}$ ,  $U_{FP}$ ,  $U_{FN}$ ,  $U_{TN}$ ). Voor de precieze gemeten waarde  $x$  geldt nu, dat de uitslag zou moeten worden beschouwd als wijzend op ziekte als:

$$\frac{P(D^+|x)}{P(D^-|x)} > \frac{U_{TN} - U_{FP}}{U_{TP} - U_{FN}}$$

en dat de uitslag zou moeten worden beschouwd als wijzend op niet-ziekte als:

$$\frac{P(D^+|x)}{P(D^-|x)} < \frac{U_{TN}-U_{FP}}{U_{TP}-U_{FN}}$$

Zijn de twee uitdrukkingen gelijk aan elkaar, dan maakt het niet uit welke van de twee conclusies men trekt.

Men kan in plaats van  $U_{TP}$ ,  $U_{FP}$ ,  $U_{TN}$  en  $U_{FN}$  beter (namelijk neutraler) schrijven respectievelijk:

$$U(B^+|D^+), U(B^+|D^-), U(B^-|D^-), U(B^-|D^+).$$

$B^+$  en  $B^-$  betekenen dan het verder behandelen van patiënten alsof ze ziekte D hebben respectievelijk niet hebben. Dit kan gebeuren in situaties waarin men in werkelijkheid wel, respectievelijk niet, de ziekte D<sup>+</sup> heeft. In deze uitdrukkingen komt beter tot uiting, dat het gewicht van een uitkomst niet éénmalig wordt vastgesteld, doch in relatie tot de verder te nemen beleidsbeslissingen gezien wordt.

Teruggaand naar figuur 3 betreffende de verdeling van hematocriet waarden bij vrouwen die wel respectievelijk niet reageren op ijzertherapie kunnen we de formule als volgt toepassen.

Stel: we vinden de waarde van de gezondheidstoestand van iemand die terecht niet behandeld wordt met ijzer ( $U(B^-|D^-)$ ) gelijk aan 100%. Tevens achten we de schade van onnodige therapie gering:  $U(B^+|D^-)$  is bijna even hoog, nl. 95%. Ook ten aanzien van de toestand van iemand die terecht met ijzer wordt behandeld zijn we positief gestemd:

$U(B^+|D^+) = 95\%$ . Het niet behandelen van iemand met ijzergebrek vinden we echter nadelig:  $U(B^-|D^+) = 60\%$ .

$$\text{We vinden nu: } \frac{U(B^-|D^-)-U(B^+|D^-)}{U(B^+|D^+)-U(B^-|D^+)} = \frac{(100-95)\%}{(95-60)\%} = \frac{1}{7}$$

$$\text{Dat wil zeggen: } \frac{P(D^+|x)}{P(D^-|x)} \text{ dient } > \frac{1}{7} \text{ te zijn,}$$

teneinde naar aanleiding van uitslag x tot zinvolle behandeling te kunnen overgaan. Dit betekent dat de gemeten hematocrietwaarde lager dan 38.5 (= x) dient te zijn (zie figuur 3).

5.3. Het bepalen van de optimale teststrategie, in relatie tot de utiliteiten van de diverse mogelijke uitkomsten.

De vraag welke normen, of beter handvaten, men moet hanteren bij de interpretatie van diagnostische gegevens kan verbreed worden naar de vraag: in welke situaties is het zinvol om een bepaalde test toe te passen? Naar aanleiding van deze vraagstelling introduceerden Pauker en Kassirer 'the threshold approach to clinical decision-making' (62). Door het uitvoeren van een diagnostische test wil men de a priori kans op een aandoening omzetten in een a posteriori kans. Het besluit om naar aanleiding van de testuitslag tot een bepaalde beleids optie (bijvoorbeeld therapie) over te gaan hangt af van de vraag of een daardoor bepaalde a posteriori kans kan worden bereikt of overschreden:  $P_x(D^+|T^+)$ .

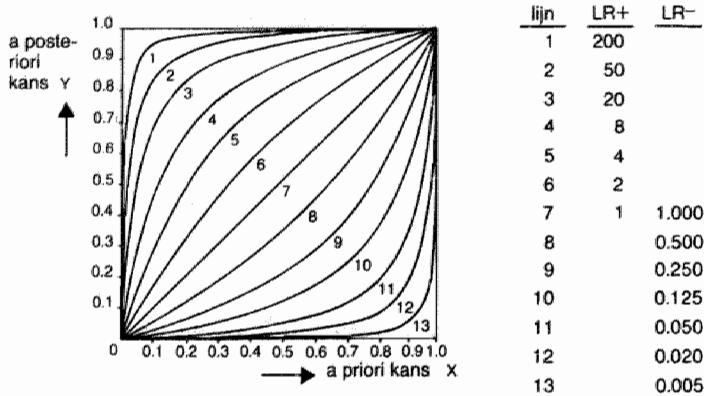
Het is duidelijk dat uitvoeren van de test geen zin heeft als:

- 1) de a priori kans op zichzelf al zoveel hoger is dan  $P_x(D^+|T^+)$  dat door het uitvoeren van de test de negatief voorspellende waarde niet lager kan komen te liggen dan  $P_x(D^+|T^+)$
- 2) de a priori kans zo laag is, dat bij een positieve testuitslag de positief voorspellende waarde niet  $P_x(D^+|T^+)$  kan bereiken.

Er is dus een bepaalde range van de a priori kans waar binnen het uitvoeren van de test zinvol is. Daarboven en daaronder dient men direct te besluiten om wel respectievelijk niet verder te behandelen. Deze range wordt begrensd door twee drempel waarden (thresholds) van de a priori kans.

Indien men van tevoren een idee heeft van de grootte van de a posteriori kans die overschreden moet worden, wil men besluiten tot behandeling of nader onderzoek, dan kan men deze drempelwaarden berekenen.

Men moet dan wel informatie hebben over de validiteit van de test. Het is mogelijk een nomogram op te stellen, met behulp waarvan deze waarden snel zijn af te lezen. In figuur 6 hebben wij dit aangegeven. Op deze wijze kan men dus een indruk krijgen van het 'indicatiegebied' voor toepassing van de test, in termen van de a priori kans. Uitgaande van de validiteit van de test (sensitiviteit en specificiteit) en de met het oog op de nadere diagnostiek te overschrijden posterior kans, kan men dit gebied gemakkelijk berekenen met behulp van een andere schrijfwijze van de formule van Bayes.



berekening van de lijnen  
doormiddel van de formules:

$$LR+ : y = \frac{1}{1 + \frac{1}{\frac{LR+ \cdot x}{1-x}}}$$

$$LR- : y = 1 - \left( \frac{1}{1 + \frac{1}{\frac{LR- \cdot x}{1-x}}} \right)$$

Figuur 6: Nomogram voor de bepaling van de a priori kansen waarbij het uitvoeren van de test zinvol is.

Toelichting:

- (1) Stel, men heeft een test beschikbaar met sensitiviteit van 56% en een specificiteit van 86%, dat wil zeggen de

$$LR+ = \frac{\text{sensitiviteit}}{1-\text{specificiteit}} = 4.00$$

$$\text{en de } LR- = \frac{1-\text{sensitiviteit}}{\text{specificiteit}} = 0.51$$

- (2) Men doet met deze test diagnostiek ten aanzien van ziekte D. Men is van mening dat behandeling van ziekte D nodig is als na het uitvoeren van de test blijkt dat de kans op de ziekte meer dan 50% is.

- (3) Men kan nu zien dat het doen van de test slechts zinvol is indien de a priori kans tussen circa 20% en circa 65% ligt. Dit zijn de snijpunten van de horizontale die de 50% a posteriori kans markeert, en respectievelijk lijn 5 (behorend bij  $LR+ = 4.0$ ) en lijn 8 (behorend bij  $LR- = 0.50$ ).

Indien de a priori kans lager is dan 20%, dan zal ook bij een positieve uitslag geen a posteriori kans van 50% bereikt worden. Indien de a priori kans hoger is dan 65%, dan zal, ook bij een negatieve uitslag, de a posteriori kans toch niet onder de 50% kunnen komen.

Tabel 8: Het indicatiegebied van de hemoculttest in termen van a priori kansen, bij verschillende drempels voor uitgebreidere diagnostiek, voor verschillende waarden van sensitiviteit en specificiteit.

		drempels voor uitgebreidere diagnostiek		
sensitiviteit	specificiteit	1% indicatie- gebied	5% indicatie- gebied	10% indicatie- gebied
67%	97%	0,05%- 2,9%	0,24%-13,4%	0,50%-24,6%
80%	83%	0,21%- 4,0%	1,10%-17,9%	2,31%-31,6%
90%	96%	0,05%- 8,8%	0,23%-33,6%	0,49%-51,6%
95%	80%	0,21%-13,9%	1,10%-45,7%	2,29%-64,0%
95%	96%	0,04%-16,2%	0,22%-50,3%	0,47%-68,1%

Toelichting:

De berekening van de grenzen van het indicatiegebied kan geschieden via de volgende schrijfwijzen van de formule van Bayes, waarbij de drempelwaarde de gewenste voorspellende waarde is.

De kans behorende bij de ondergrens =

$$\frac{1}{1 + \left( \frac{\text{sensitiviteit}}{100\% - \text{specificiteit}} \right) \times \left( \frac{100\% - \text{drempelwaarde}}{\text{drempelwaarde}} \right)}$$

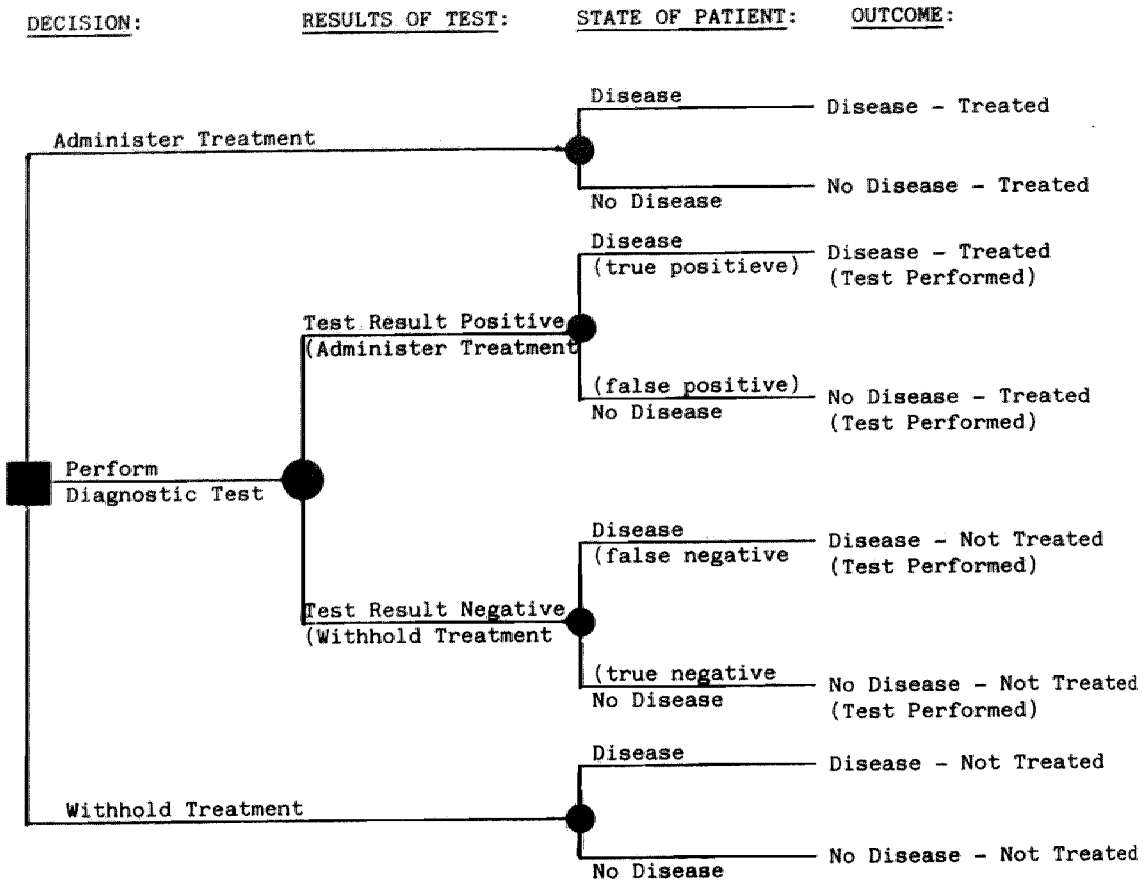
De kans behorende bij de bovengrens =

$$\frac{1}{1 + \left( \frac{100\% - \text{sensitiviteit}}{\text{specificiteit}} \right) \times \left( \frac{100\% - \text{drempelwaarde}}{\text{drempelwaarde}} \right)}$$

Als voorbeeld geven wij in tabel 8 de berekening van het indicatiegebied voor de hemocult test (op occult bloed in de faeces ter opsporing van dikke darm kanker) voor diverse 'drempel kansen' voor nadere diagnostiek (colonfoto en/of coloscopie) en verschillende op grond van de literatuur te beargumenteren waarden van sensitiviteit en specificiteit (deze laatste kunnen variëren met de wijze van uitvoering).

Indien men een opvatting heeft over de waarde (utiliteiten) van de diverse testuitkomsten kan men een stapje verder gaan. De situatie kan dan weer in kaart worden gebracht door middel van een beslisboom (figuur 7).





Figuur 7: Beslisboom, ter bepaling van de drempelwaarden van de a priori kans om het toepassen van een test zinvol te doen zijn (62).

Het is duidelijk dat de drempelwaarden  $P_1(D^+)$  en  $P_2(D^+)$  afhankelijk zijn van het nut van wel en niet behandelen bij zieken en niet-zieken, de risico's van de test en de validiteit van de test. Gedefiniëerd worden:

- . de utiliteit van verdere behandeling bij zieken:  $U(B^+|D^+)$
- . de utiliteit van verdere behandeling bij niet-zieken:  $U(B^+|D^-)$
- . de utiliteit van geen behandeling bij zieken:  $U(B^-|D^+)$
- . de utiliteit van geen behandeling bij niet-zieken:  $U(B^-|D^-)$

- . de risico's van het uitvoeren van de test (inclusief de mate van belasting van de test):  $R_t$
- .  $P(T^+|D^+)$  ( $= 1-P(T^-|D^+)$ ): de sensitiviteit van de test
- .  $P(T^-|D^-)$  ( $= 1-P(T^+|D^-)$ ): de specificiteit van de test
- .  $P(D^+)$  : de a priori kans op D.

De verwachte utiliteit (EU) van de mogelijke diagnostische strategie is nu uit te drukken met behulp van deze symbolen.

(1) EU (niet testen + niet behandelen) =

$$P(D^+) \cdot U(B^-|D^+) + [1-P(D^+)] \cdot U(B^-|D^-)$$

(2) EU (testen) =

$$P(T^+|D^+) \cdot P(D^+) \cdot [U(B^+|D^+) - R_t] + P(T^-|D^+) \cdot P(D^+) \cdot [U(B^-|D^+) - R_t] + \\ P(T^+|D^-) \cdot P(D^-) \cdot [U(B^+|D^-) - R_t] + P(T^-|D^-) \cdot P(D^-) \cdot [U(B^-|D^-) - R_t]$$

en

(3) EU (niet testen behandelen) =

$$P(D^+) \cdot U(B^+|D^+) + [1-P(D^+)] \cdot U(B^+|D^-).$$

Hierbij is uitgegaan van de (vereenvoudigde) situatie, dat  $R_t$  niet afhankelijk is van de ziekte status.

Voor  $P_1(D^+)$ , d.w.z. de drempel tussen strategie 1 (niet testen en niet behandelen) en strategie 2 (testen en afhankelijk van de uitslag handelen) geldt dat verwachte utiliteiten van beide strategieën juist aan elkaar gelijk te zijn. Door (1) en (2) aan elkaar gelijk te stellen kan men de vergelijking dan oplossen voor  $P(D^+)$ :

$$P_1(D^+) =$$

$$\frac{P(T^+|D^-) \cdot [U(B^-|D^-) - U(B^+|D^-)] + R_t}{P(T^+|D^-) \cdot [U(B^-|D^-) - U(B^+|D^-)] + P(T^+|D^+) \cdot [U(B^+|D^+) - U(B^-|D^+)]}$$

Evenzo kan men  $P_2(D^+)$  berekenen door (2) en (3) aan elkaar gelijk te stellen:

$$P_2(D^+) =$$

$$\frac{P(T^-|D^-) \cdot [U(B^-|D^-) - U(B^+|D^-)] - R_t}{P(T^-|D^-) \cdot [U(B^-|D^-) - U(B^+|D^-)] + P(T^+|D^+) \cdot [U(B^+|D^+) - U(B^-|D^+)]}$$

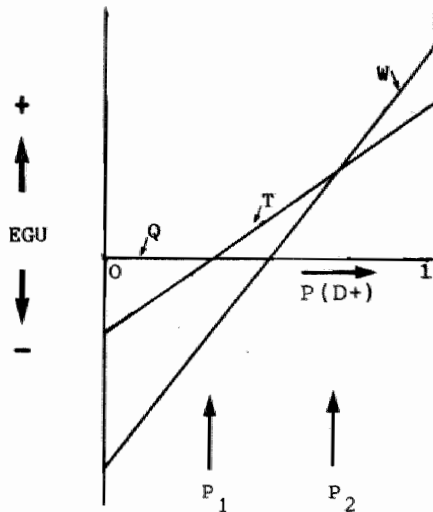
Indien men de testkenmerken, de waarden van de diverse uitkomsten, en de risico's van de test kent of kan schatten, kan men het verloop van de verwachte utiliteit van de drie strategieën (ten opzichte van elkaar) grafisch weergeven als functies van  $P(D^+)$  (figuur 8) (63). De snijpunten tussen de grafieken van Q en T respectievelijk T en W leveren de drempelwaarden  $P_1(D^+)$  en  $P_2(D^+)$  op. Men heeft dan in feite een sensitiviteitsanalyse verricht, waarbij wordt nagegaan hoe de voorkeur voor een bepaalde strategie verandert bij variatie van  $P(D^+)$ . Deze benadering beperkt zich niet tot testen met dichotome (of gedichotomiseerde) uitkomsten. Doubilet (63) heeft de benadering uitgebreid voor tests met een ordinale, respectievelijk een continue testvariabele. In zijn benadering is het mogelijk de optimale strategie te kiezen uitgaande van een bepaalde a priori kans én iedere gemeten testuitkomst  $T_x$ . Het is mogelijk om daarbij een keuze te maken tussen diverse tests of teststrategieën.

In feite komt het erop neer, dat het kiezen van de meest optimale diagnostische strategie een besliskundige probleemstelling is, waarbij naast de structuur van de te nemen beslissing (te representeren in een beslisboom) a priori kansen, conditionele kansen en utiliteiten de pijlers van de optimale besluitvorming zijn.

### Conclusie.

In dit hoofdstuk hebben wij een inventarisatie gegeven van de vragen die aan de orde komen bij het spreken over referentie waarden, en de diverse benaderingen die er zijn voor de beantwoording ervan.

De belangrijkste conclusie is, dat het definiëren van een referentiegebied c.q. optimaal 'afkappunt' een probleem is dat niet in het algemeen is op te lossen. De a priori kans, de testkenmerken (c.q. de verdeling van uitslagen bij zieken en niet-zieken) de risico's van de test, en het gewicht van de diverse uitkomsten spelen een grote rol.



Figuur 8 Grafische weergave van de bepaling van de drempelwaarden van de a priori kans voor het zinvol zijn van het verrichten van een diagnostische test (63).

Toelichting:

- Q = niet testen, niet verder behandelen ("niets doen")  
 T = testen en naar aanleiding van de uitslag handelen ("testen")  
 W = niet testen, direct verder handelen ("direct behandelen")  
 EGU = expected gain in utility  
 $P_1$  = a priori kans, die de drempel vormt tussen "niets doen" en "testen"  
 $P_2$  = a priori kans, die de drempel vormt tussen "testen" en "direct behandelen".

In klinische termen: de patiënt, diens probleem en persoonlijke waarderungen van risico's en uitkomsten moeten zo individueel mogelijk benaderd worden. In feite heeft men te maken met een beslistkundige probleemstelling, waarbij ook de zin van het doen van de test ter discussie staat.

Diagnostische informatie die in principe wel in het algemeen te verschaffen is, betreft de verdelingen van test-uitkomsten bij zieken en niet-zieken, liefst voor verschillende ziekte-toestanden en verschillende populaties. Deze vormen immers de sleutel voor de testkenmerken, welk afkappunt of methode men ook kiest, en zijn de basis voor iedere

benadering. De voorkeur verdient hierbij een benadering waarbij, in geval van ordinale of continu-verdeelde testuitkomsten, zoveel mogelijk de informatie die vervat is in de precieze waarde van de testuitslag wordt gebruikt in plaats van alle uitkomsten onder te brengen in een kunstmatig gedichotomiseerde schaal.

Een belangrijk uitgangspunt is het doel dat men nastreeft. Of het nu gaat om het beschrijven van de verdeling van meetwaarden in populaties; het bestuderen van het verband tussen testuitslagen, klachten en diagnoses; het selecteren van tests met een groot onderscheidend vermogen; medische besluitvorming met een behoefte aan toename aan zekerheid; prognostische voorspellingen, of om het bepalen van de optimale test-strategie in individuele gevallen; telkens is een op het doel toegesneden benaderingswijze aangewezen.

#### LITERATUUR.

1. Murphy EA. The normal, and the perils of the sylleptic argument. *Perspectives in Biology and Medicine* 1972; 566-581.
2. Meyers LD, Habicht JP, Johnson CC. Components of the differences in hemoglobin concentrations in blood between black and white women in the United States. *Am J Epid* 1979; 109: 535-549.
3. Owen GM, Yanochik-Owen A. Should there be a different definition of anemia in black and white children? *Am J Public Health* 1977; 67: 865-866.
4. Micozzi M. On definition of anemia in pregnancy. *Am J Public Health* 1978; 68: 907-908.
5. Wintrobe MM. *Clinical Haematology*. Lea & Febiger 1981.
6. Kilpatrick GS. Prevalence of anaemia in the general population. *Brit Med J* 1961; dec 30: 1736-1738.
7. Waters WE. Ten-year Haematological Follow-up: mortality and haematological changes. *British Medical Journal* 1969; 4: 761-764.
8. Elwood PC, Waters WE, Greene WJ, Wood MM. Evaluation of a screening survey in adult non-pregnant women. *Brit Med J* 1967; 4: 714-717.
9. Garby L, Imnell L, Werner I. *Acta Soc Med Upsalien* 1967; 72: 91.

10. Hallberg L, Bengtsson C, Garby L, Tibblin E, Lennartson J, Rossander L. An analysis of factors leading to a reduction in iron deficiency in Swedish women. *Bulletin WHO* 1979; 57: 947-954.
11. Garby L, Imell L, Werner I. Iron deficiency in women of fertile age in a Swedish community. III. Estimation of prevalence based on response to iron supplementation. *Acta Med Scand* 1969; 185: 113-117.
12. Elwood PC. Nutritional state of elderly Asian and English subjects in Coventry. *Lancet* 1972; 3 juni: 1224-1227.
13. Bengtsson C. The study of women in Gothenburg 1968-1969 - a population study. *Acta Med Scand* 1973; 193: 311-318.
14. Bengtsson C. The population study of women in Gothenburg 1974-1975 - the second phase of a longitudinal study. *Scand J Soc Med* 1978; 6: 49-54.
15. Campbell H. Pilot survey of haemoglobin and plasma urea concentration in a random sample of adults in Wales 1965-1966. *Brit J Prev Soc Med* 1968; 22: 41-49.
16. De Wijn JF, Rusbach HW. Het hemoglobinegehalte van Nederlandse schoolkinderen en adolescenten. *Ned Tijdschr Geneesk* 1961; 105: 1028-1035.
17. Nijenhuis HWA. Pre-employment health conditions among young people. *Proefschrift, Amsterdam* 1975.
18. Hoeven J van der. Evaluation of a programme of periodic medical examination. *Proefschrift, Nijmegen* 1971.
19. International Committee for Standardization in Haematology. Recommendation for measurement of erythrocyte sedimentation rate of human blood. *Am J Clin Pathol* 1977; : 505-507.
20. Lewis, SM. Erythrocyte Sedimentation Rate and plasma viscosity. *Association of Clinical Pathologists Broadsheet* 1980; 94: 1-7.
21. Böttiger LE, Svedberg CA. Normal erythrocyte sedimentation rate and age. *Br Med J* 1967; 85-87.
22. Raffnson V, Bengtsson C, Lennartsson J, Noppa H, Tibblin E. Erythrocyte sedimentation rate in a population sample of women with special reference to its clinical and prognostic significance. *Acta Med Scand* 1979; 206: 207-214.
23. Miller A, Green M, Robinson D. Simple rule for calculating normal erythrocyte sedimentation rate. *Br Med J* 1983; 286: 266.

24. Eastham RD. Simple rule for calculating normal erythrocyte sedimentation rate (letter to the editor). Brit Med J 1983; 22 jan: 557-558.
25. Patri B, Margara JP, Loison F, Dubrisay J. Vitesse de sédimentation. Le Concours Medical 1980; 102: 4265-4269.
26. Costongs GMPJ. Intra-individual variations and critical differences of clinical laboratory parameters. Proefschrift. Maastricht 1985.
27. Feinstein AR. Clinical Epidemiology, the architecture of clinical research. WB Saunders, Philadelphia
28. Knottnerus JA, Knipschild PG, Sturmans F. Symptomatie en selectiebias, vertekening van het verband tussen klachten en diagnosen ten gevolge van selectie naar hogere echelons. Huisarts en Wetenschap 1985; 28: 325-330.
29. Knottnerus JA, Knipschild PG, Wersch JWJ van, Sijstermanns AHJ. Onverklaarde moeheid en hemoglobinegehalte, een onderzoek vanuit de huisartspraktijk. Ned Tijdschr Geneesk 1986; 130: 402-405.
30. Knipschild PG, Visser A, Knottnerus JA. Hoofdpijn en hoge bloeddruk van een bevolkingsonderzoek in Amsterdam. Huisarts en Wetenschap 1984; 27: 441-445.
31. Knottnerus JA, Knipschild PG, Wersch JWJ van, Sijstermanns AHJ. Serumferritine, bij moeheidspatiënten en gezonde controle personen in de huisartspraktijk. Wordt gepubliceerd in: Ned Tijdschr Geneesk 1986; 130.
32. Bezemer PD, Netelenbos JC, Mulder C, Theune JA, Stamhuis IH, Straub JP. Determining reference ('normal') limits in medicine: an application. Statistics In Medicine 1983; 2: 191-198.
33. Pincherle G, Shanks J. Value of the erythrocyte sedimentation rate as a screening test. Br J Prev Soc Med 1967; 21: 133-136.
34. Boyd RV, Hoffbrand BI. Erythrocyte sedimentation rate in elderly hospital in-patients. Br Med J 1966; : 901-902.
35. Rafnsson V, Bengtsson C, Lurie M. Erythrocyte Sedimentation Rate in women with different manifestations of joint disease. Scand J Rheumatol 1982; 11: 87-95.
36. Kenny MW, Worthington DJ, Stuart J, AJ Davies, M Farr, Davey PG, Chughtai MA. Efficiency of haematological screening tests for detecting disease. Clin Lab Haematol 1981; 3: 299-305.

37. Knuth UA, Mühlenstedt D, Schneider HPG. Zur effektivität von Kurszwellenbelastungstests, Blutkörperchen senkungsgeschwindigkeit und leucocytenzahl bei der diagnose der subakuten adnexitis. Geburtshilfe und Frauenheilkunde 1982; 42: 52-55.
38. Peyman MA. The effect of malignant disease on the erythrocyte sedimentation rate. Br J Cancer 1962; 16: 56-71.
39. Rai GS. Erythrocyte sedimentation rate and disease in the elderly. J Am Geriatr Soc 1979; 27: 382-383.
40. Asbjörnsen G. Serum copper compared to erythrocyte sedimentation rate as indicator of disease activity in Hodgkin's disease. Scand J Haemat 1977; 22: 193-196.
41. Wilimas J, Thompson E, Smith KL. Value of serum copper levels and erythrocyte sedimentation rates as indicators of disease activity in children with Hodgkin's disease. Cancer 1978; 42: 1929-1935.
42. Salter RH. Gastrointestinal malignancy and the ESR. Practitioner 1981; 255: 566-568.
43. Schwartz WB, Wolfe HJ, Pauker SG. Pathology and probabilities, a new method to interpreting and reporting biopsies. New Engl J Med 1981; 305: 917- 923.
44. Kottnerus JA, Sturmans F. Over vertekening van testuitslagen door reeds aanwezige informatie. The Practitioner 1985; 2: 287-295.
45. Sturmans F. Epidemiologie, theorie, methoden en toepassing. Nijmegen 1982. Dekker & Van der Vegt.
46. Begg CB, Greenes RA. Assessment of diagnostic tests when disease verification is subject to selection bias. Biometrics 1983; 39: 207-215.
47. Gray R, Begg CB, Greenes RA. Construction of Receiver Operating Characteristics when disease verification is subject to selection bias. Med Decis Making 1984; 4:151-164.
48. Weinstein MC, Fineberg HV. Clinical Decision Analysis. Saunders, Philadelphia 1980.
49. Department of Clinical Epidemiology & Biostatistics. McMaster University. Manual for a workshop on how to teach the critical appraisal of clinical evidence (Diagnosis Unit). Hamilton 1984.
50. Conell FA, Koepsell TD. Measures of gain in certainty from a diagnostic test. Am J Epid 1985; 121: 744-753.



51. McNeill BJ, Keeler E, Adelstein SJ. Primer on certain elements of medical decision making. *N Eng J Med* 1975; 293: 211-215.
52. Dawber TR. The Framingham Study. Harvard University Press Cambridge, 1980.
53. Herman B, Leyten ACM, Luijck JH van, Frenken GWGM, Op de Coul AAW, Schulte BPM. An evaluation of risk factors for stroke in a Dutch Community. *Stroke* 1982; 13: 334-339.
54. Tibblin G, Wilhelmsen L, Werkö L. Risk factors for myocardial infarction and death due to ischemic heart disease and other causes. *Am J Cardiology* 1975; 35: 514-522.
55. Yano K, Reed DM, McGee DL. Ten-year incidence of coronary heart disease in the Honolulu heart program. *Am J Epidemiol* 1984; 119: 653-666.
56. Elwood PC, Waters WE, Benjamin IT, Sweetnam PM. Mortality and anaemia in women. *The Lancet* 1974; 2: 891-894.
57. Campbell MJ, Elwood PC, Mackean J, Waters WE. Mortality, haemoglobin level and haematocrit in women. *J Chron Dis* 38: 881-889.
58. Truett J, Cornfield J, Kannell W. A multivariate analysis of coronary heart disease in Framingham. *J Chron Dis* 1967; 20: 511-524.
59. Abu-Zeid HAH. The relation between haemoglobin level and the risk for ischemic heart disease: a prospective study. *J Chron Dis* 1976; 29: 395-403.
60. Wilhelmsen L, Wedel H, Tibblin G. Multivariate analysis of risk factors for coronary heart disease. *Circulation* 1973; 48: 950-958.
61. Brand RJ, Rosenman RH, Sholtz RI, Friedman M. Multivariate prediction of coronary heart disease in the Western Collaborative Group Study compared to the findings of the Framingham Study. *Circulation* 1976; 53: 348-355.
62. Pauker SG, Kassirer JP. The threshold approach to clinical decision making. *New Eng J Med* 1980; 302: 1109-1117.
63. Doubilet P. A mathematical approach to interpretation and selection of diagnostic tests. *Med Decis Making* 1983; 3: 177-195.

---

DE INVLOED VAN VERWIJSGEDRAG OP HET ONDERSCHIEDEND VERMOGEN VAN DIAGNOSTISCHE TESTS.

J.A. Knottnerus <sup>1)</sup> - P. Leffers <sup>2)</sup>

Samenvatting.

Het onderscheidend vermogen van een diagnostische test, gekarakteriseerd door sensitiviteit, specificiteit, aannemelijkheidsquotient en ROC-curve kan veranderen ten gevolge van verwijspatronen van huisartsen. Vooral de symptomatologie en de de uitslag van de eventueel reeds verrichte test zullen de verwijskans beïnvloeden, terwijl de mate van ontwikkeling van het pathologisch proces rechtstreeks de kans op positieve uitslagen beïnvloedt. De mogelijke veranderingen worden geanalyseerd met behulp van rekenvoorbeelden, waarbij achtereenvolgens worden bestudeerd: de verwijzing is alleen afhankelijk van de symptomatologie; de verwijzing is afhankelijk van symptomatologie en testuitslag; de verwijzing is alleen afhankelijk van de testuitslag. In het eerste geval blijven de testkenmerken en voorspellende waarden per stratum van de symptomatologie onveranderd, in het laatste geval veranderen alleen de voorspellende waarden niet. Als er een positief verband bestaat tussen de afwijkende testuitslag en de verwijskans, zal overall de sensitiviteit toenemen, de specificiteit afnemen en het aannemelijkheidsquotient van de afwijkende uitslag afnemen. Een algemene representatie wordt gegeven voor de evaluatie van de richting van de de verandering van het aannemelijkheidsquotient.

1) Capaciteitsgroep Huisartsgeneeskunde en Capaciteitsgroep Epidemiologie

2) Capaciteitsgroep Epidemiologie

Rijksuniversiteit Limburg

### Inleiding.

Vaak wordt aangenomen, dat de validiteit van een diagnostische test (in casu de sensitiviteit en specificiteit) een vast gegeven is, onafhankelijk van de populatie waarin men de test toepast. De voorspellende waarde van testuitslagen wordt dan verder alleen bepaald door de prevalentie van de te diagnostiseren ziekte in de te onderzoeken populatie. Naar aanleiding hiervan is erop gewezen dat testuitslagen in de huisartspraktijk een andere voorspellende waarde kunnen hebben dan in de praktijk van de specialist. De prevalenties (a priori kansen) zullen namelijk in beide situaties dikwijls verschillen (1-4).

In de laatste jaren is benadrukt, dat niet alleen de voorspellende waarde, maar ook de validiteit van de test zelf niet los te zien is van de populatie die onderzocht wordt (3, 5-8). Behalve de aanwezigheid of afwezigheid van ziekte (tot uiting komend in de prevalentie) zijn er andere factoren die de testuitslag kunnen beïnvloeden, en deze factoren kunnen in verschillende populaties anders verdeeld zijn.

Zo is beschreven, dat de sensitiviteit van reumaserologie ten aanzien van reumatoïde artritis in de algemene bevolking lager is, dan in, door middel van verwijzing naar reumapoliklinieken, geselecteerde populaties (9). En wat betreft de mammografische screening op borstkanker wordt in de algemene bevolking vaak een hogere specificiteit gevonden dan bij vrouwen die voor evaluatie van een knobbeltje naar een chirurgisch centrum zijn verwezen (10). Met dergelijke verschillen moet men rekening houden bij medisch-besliskundige analyses en het op grond daarvan aanbevelen van bepaalde handelwijzen. Vertekeningen, en dus foutieve adviezen kunnen het gevolg zijn van onjuiste schattingen van de validiteit van diagnostische tests.

Vaak zal het niet mogelijk zijn om de benodigde basisinformatie rechtstreeks uit de doelpopulatie te betrekken. Het is bijvoorbeeld nauwelijks denkbaar dat de validiteit van het inspannings-electrocardiogram bij mensen die de huisarts bezoeken met atypische klachten op de borstcoit zal worden vastgesteld door bij een ruime steekproef van deze personen, naast electrocardiografie, ook nog coronairangiografie te verrichten (11). In dergelijke gevallen zullen gegevens ontleend aan meer of minder geselecteerde groepen het uitgangspunt blijven. Om die

reden is het van belang om tenminste inzicht te hebben in de mechanismen die bij de selectie een rol spelen, teneinde uitspraken te kunnen doen over de te verwachten richting en grootte van de vertekening. Selectie-invloeden kunnen een vertekening geven van het onderscheidend vermogen, indien we te maken hebben met een "niet-perfecte test", wanneer er dus een overlap is van uitslagen van zieken en niet-zieken. Dit is in de regel het geval.

Dit artikel is gericht op verschillen die te verwachten zijn tussen de huisartsgeneeskundige en de specialistische praktijk, wat betreft de validiteit - of liever: het onderscheidend vermogen - van diagnostische tests. Besproken wordt het effect van selectie door middel van verwijzing op diverse parameters waarin men het onderscheidend vermogen kan uitdrukken:

- de sensitiviteit en de specificiteit.
- het aannemelijkheidsquotient.

$$\text{. van de positieve uitslag} = \frac{\text{sensitiviteit}}{1 - \text{specificiteit}}$$

$$\text{. van de negatieve uitslag} = \frac{1 - \text{sensitiviteit}}{\text{specificiteit}}$$

- de Receiver Operating Characteristic (ROC-curve).

Wij beginnen met een bespreking van enkele factoren die in het kader van het selectieproces belangrijk zijn.

#### Selectieve invloeden bij verwijzing.

Verandering van het onderscheidend vermogen van een diagnostische test door verwijzing kan optreden indien er factoren werkzaam zijn, die zowel met de kans op bepaalde (bijvoorbeeld positieve) testuitslagen als met de kans op verwijzing samenhangen. Zulke factoren zijn:

- a) het "klinisch beeld": de ontwikkeling van de symptomatologie en ernst van de ziekte;
- b) de uitslagen van reeds verricht diagnostisch onderzoek.

- . Symptomatologie kan in allerlei gradaties van ontwikkeling voorkomen. Een aandoening kan gepaard gaan met een breed spectrum van klinische beelden, variërend van "asymptomatisch", tot "sterk ontwikkeld", of "zeer suspect". Ook bij mensen zonder de te diagnostiseren aandoening kan een meer of minder sterk ontwikkeld beeld voorkomen. Er kan bijvoorbeeld een palpabele tumor in de mamma aanwezig zijn zonder dat er sprake is van kanker.
- . Wanneer er sprake is van ziekte, zijn er weer diverse graden van ernst (hier is bedoeld: de ontwikkeling van het pathologisch proces) te onderscheiden. Dit is bijvoorbeeld het geval bij de histologische en klinische ontwikkelings-stadia van veel maligne tumoren. Hoewel zij in geval van ziekte in het algemeen positief samenhangen, dienen de mate van ontwikkeling van de symptomatologie (S) en de ernst van ziekte (D) van elkaar te worden onderscheiden. Immers een ziekte in een vergevorderd stadium kan nog asymptomatisch zijn, terwijl een niet- of niet-ernstig zieke toch een verontrustende symptomatologie kan vertonen. Er zijn ook situaties waarbij S en D geheel of grotendeels samenvallen, zoals bij allerlei "symptoomdiagnosen" (hoofdpijn, migraine, low back pain, angina pectoris) en in het geval dat de (ernst van de) ziekte (mede) wordt gedefinieerd aan de hand van de symptomatologie (reumatoïde artritis, coronair aandoeningen).
- . Een diagnostische test ontleent zijn bestaansrecht aan het feit dat de uitslag ervan, en daarmee de kans dat de referentiewaarde wordt overschreden (c.q. de test positief uitvalt), afhangt van de aanwezigheid van (een) bepaalde ziekte(n). De uitslag is in het algemeen ook afhankelijk van de ernst van de ziekte. Bij ernstige of verder ontwikkelde hepatitis zullen de uitslagen gemiddeld hoger uitvallen dan bij lichte hepatitis. En grote levertumoren zullen eerder een afwijkend echogram opleveren dan kleine.  
De uitslag kan voorts samenhangen met de symptomatologie. Voor een deel is dit het gevolg van het verband tussen symptomatologie en de ernst van de ziekte. Deels echter kan de symptomatologie een zelfstandige invloed hebben, zowel bij personen met een bepaalde ziekte als bij personen zonder die ziekte. Soms hangt de symptomatologie dan samen met andere afwijkingen. Zo zijn er bij vrouwen met gave

mmae minder fout-positieve mammogrammen ten aanzien van carcinoom te verwachten, dan bij vrouwen met palpabele multipiele (benigne) tumoren. Ook kan het zijn, dat een maligne tumor minder gemakkelijk wordt gevonden als hij te midden van goedaardige aandoeningen ligt (12-14). De beïnvloedingen van de kans op een bepaald test resultaat zijn complex, en de richting waarin de invloed werkt is niet altijd voorspelbaar.

. De kans dat een patiënt door de huisarts wordt verwezen naar een specialist, wordt onder meer beïnvloed door:

- 1) De mate van ontwikkeling van de symptomatologie S (inclusief voorgeschiedenis en anamnese). Met name als de ziektestatus nog onbekend is, zal de symptomatologie een belangrijke rechtstreekse invloed op de verwijskans hebben.
- 2) De ernst van de ziekte. Deze zal, los van de symptomatologie, alleen invloed op de verwijskans hebben als de huisarts naast de symptomen nog andere (terechte) gronden heeft om aan (een bepaalde graad van ernst van) ziekte D te denken én bij ernstiger vormen van ziekte D relatief vaak verwijzing wenselijk acht. Ook kan het zijn dat de ziekte reeds door de huisarts is vastgesteld. Verwijzing kan dan om therapeutische redenen plaatsvinden, vooral bij ernstiger gevallen.
- 3) De uitslag van het eventueel door de huisarts verrichte onderzoek T. Een afwijkende uitslag zal in de regel een positieve invloed op de verwijskans hebben, een niet-afwijkende uitslag een negatieve. Ook het tegengestelde is mogelijk, afhankelijk van de zin van verwijzing bij het vermoeden of vastgesteld zijn van (een) bepaalde ziekte(n). Daarnaast zal een aantal patiënten onderzoek T niet ondergaan, maar op andere gronden verwezen worden. De uitslag van de test heeft in die gevallen geen rechtstreekse invloed op de verwijskans.

In het algemeen is de kans dat de test überhaupt wordt uitgevoerd afhankelijk van de symptomatologie. Dit leidt er toe dat van de eventuele afwijkende uitslagen er relatief veel in de meer suspecte groepen worden gevonden, zelfs als er in werkelijkheid géén verband is tussen de testuitslag en de ziekte. Als nu de verwijskans toeneemt bij meer suspecte symptomatologie en bij een posi-

tieve testuitslag, dan zal dit invloed hebben op de validiteit van de test in de via verwijzing geselecteerde populatie. Bij het bestuderen van de invloed van verwijzing op het onderscheidend vermogen van test T, is de verwijskans op grond van de test zelf van belang als de test-uitslagen vóór en ná verwijzing met elkaar samenhangen. Dit is in de regel het geval, en wel via een positieve samenhang.

Ter vereenvoudiging gaan wij er in het volgende vanuit dat een bepaalde testuitslag vóór en ná verwijzing hetzelfde zal uitvallen. Het gaat dan in de verdere bespreking uiteraard om diagnostische tests, die zowel door de huisarts als de specialist uitgevoerd kunnen worden. Het zijn juist dergelijke tests waarvoor het van belang is de invloed van selectie op het onderscheidend vermogen te bestuderen.

#### Verwijzing alleen naar aanleiding van de symptomen.

Aangenomen mag worden dat, indien de diagnose nog niet bekend is, de symptomatologie (bijvoorbeeld de hevigheid van angina pectoris klachten) een rechtstreekse invloed heeft op de verwijskans. De mate van ontwikkeling van het pathologisch proces (bijvoorbeeld de graad van vernauwing van de coronair vaten) heeft in deze situatie vooral een rechtstreekse invloed op de testuitslag (de kans op een afwijkend inspanningselectrocardiogram).

Ten aanzien van niet-zieken kan men twee situaties onderscheiden:

- 1) de kans van niet-zieken op een afwijkende uitslag is onafhankelijk van de symptomatologie. Bijvoorbeeld: het inspanningselectrocardiogram van een overigens gezonde man met oesophagus-klachten die op angina pectoris lijken, wordt niet beïnvloed door de hevigheid van die klachten;
- 2) de kans van niet-zieken op een afwijkende uitslag hangt wél samen met de symptomatologie. Wij noemden in dit verband al de samenhang tussen symptomatologie van (goedaardige) mamma-afwijkingen en de validiteit van mammografie. Ook is het mogelijk dat de testuitslag vertekend is door voorafkennis over de symptomatologie (15).

SCHEMA 1: Invloed van verwijzing op de parameters betreffende het verband tussen testuitslag en ziekte, indien er een positieve samenhang is tussen de kans op een positieve testuitslag en de mate van ontwikkeling van de symptomatologie, en indien de verwijskans V groter wordt bij meer ontwikkelde symptomatologie.

Een getallenvoorbeeld.

(a) Verdeling van personen, verwijskanssen, en de kans op positieve uitslagen naar symptomatologie en ziekte-status.

Symptomatologie	Ziek	Niet-ziek
duidelijk ontwikkeld ( $S_1$ )	$P_1(T+ D+) = 0.95$ $N_1 = 40$ $V_1 = 0.8$	$P_1(T+ D-) = 0.30$ $N_1 = 40$ $V_3 = 0.8$
minder duidelijk ontwikkeld ( $S_2$ )	$P_2(T+ D+) = 0.60$ $N_2 = 60$ $V_2 = 0.2$	$P_2(T+ D-) = 0.05$ $N_2 = 160$ $V_4 = 0.2$

(b) De parameters van de testuitslag in relatie tot de ziekte vóór en ná verwijzing.

Parameters	vóór verwijzing 0.33	ná verwijzing 0.41
prevalentie		
sensitiviteit	$\frac{\sum \{Ni \times Pi(T+ D+)\}}{\sum Ni} = 0.74$	$\frac{\sum \{Ni \times Vi \times Pi(T+ D+)\}}{\sum \{Ni \times Vi\}} = 0.85$
specitiviteit	$\frac{\sum \{Ni \times Pi(T- D-)\}}{\sum Ni} = 0.90$	$\frac{\sum \{Ni \times Vi \times Pi(T- D-)\}}{\sum \{Ni \times Vi\}} = 0.83$
Aannemelijks- heidsquotient T		
positieve uitslag	7,4	5,0
negatieve uitslag	0,26	0,18
Voorspellende waarde T		
overall:		
positieve uitslag	0,79	0,77
negatieve uitslag	0,87	0,89
S1:		
positieve uitslag	0,76	0,76
negatieve uitslag	0,93	0,93
S2:		
positieve uitslag	0,82	0,82
negatieve uitslag	0,86	0,86

In het voorbeeld in schema 1 zijn wij uitgegaan van een duidelijk en een minder duidelijk ontwikkelde symptomatologie, en een positieve samenhang tussen mate van ontwikkeling van de symptomatologie en de



ernst van de ziekte. Dit levert een hogere sensitiviteit van de test op bij meer ontwikkelde symptomatologie. Ook voor de niet-zieken in dit voorbeeld geldt dat de kans op positieve uitslagen (100%-specificiteit) groter wordt bij meer ontwikkelde symptomen zoals bijvoorbeeld het geval is bij mammografie in aanwezigheid van benigne tumoren. Als men nu het onderscheidend vermogen van de test evalueert in de door middel van verwijzing geselecteerde populatie, en daarbij uitgaat van een hogere verwijskans bij meer ontwikkelde symptomatologie, dan blijkt onder meer dat:

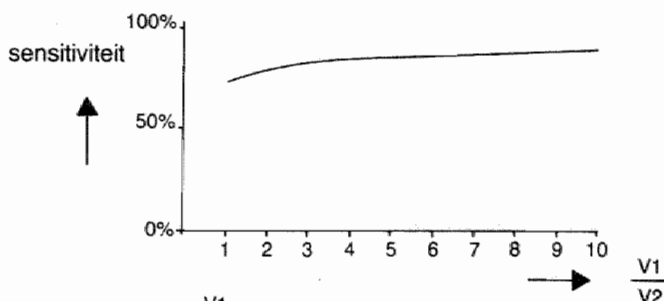
- 1) de overall-sensitiviteit duidelijk toeneemt;
- 2) de overall-specificiteit afneemt;
- 3) het aannemelijkheidsquotiënt van zowel de positieve als de negatieve uitslag afnemen.

Ook de andere parameters - de prevalentie en de voorspellende waarden - veranderen. Zoals men verwacht is de prevalentie na verwijzing hoger, maar daarentegen vindt men in dit voorbeeld een lagere voorspellende waarde van de positieve uitslag en een hogere van de negatieve uitslag! De reden hiervoor is dat bij de niet-zieken met een meer ontwikkelde symptomatologie de specificiteit lager is en de verwijskans hoger. Indien zoals in het voorbeeld de verwijskans voor zieken en niet-zieken bij ernstiger symptomatologie met dezelfde factor toeneemt (namelijk  $V1/V2 = V3/V4$ ), dan is het verband tussen deze factor en de genoemde testkenmerken grafisch eenvoudig weer te geven (figuur 1).

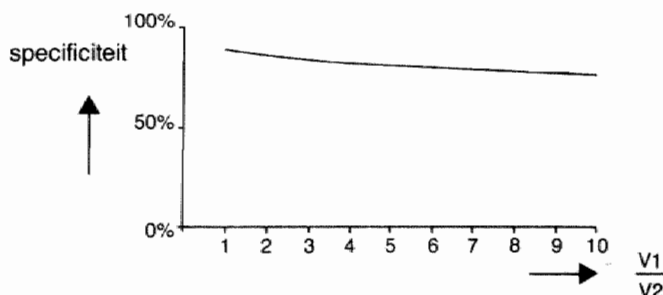
Indien de kans op een positieve uitslag bij niet-zieken niet afhankelijk is van de symptomatologie, zal de specificiteit bij verwijzing alleen op geleide van de symptomatologie niet veranderen. Het aannemelijkheidsquotiënt van een positieve test neemt dan evenredig toe met de sensitiviteit.

Het veelbeschreven patroon, dat door verwijzing de prevalentie en de voorspellende waarden van een positieve testuitslag toenemen, wordt in ons voorbeeld maar ten dele bevestigd. De in het algemeen gevolgde redenering dat men de voorspellende waarde van een test, na voorselectie via verwijzing, kan berekenen uit de hogere prevalentie bij gelijkblijvende testeigenschappen lijkt te simpel.

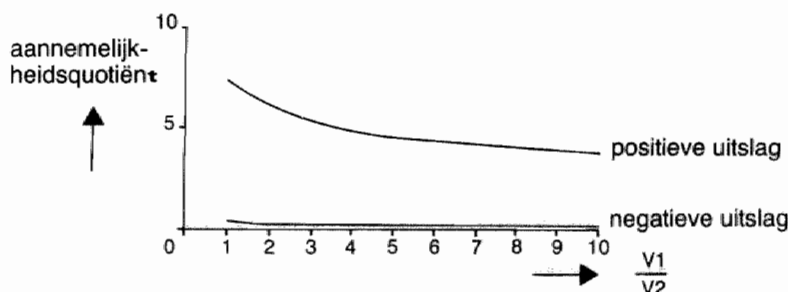
**A** verband tussen  $\frac{V_1}{V_2}$  en de sensitiviteit



**B** verband tussen  $\frac{V_1}{V_2}$  en de specificiteit



**C** verband tussen  $\frac{V_1}{V_2}$  en de aannemelijkheidsquotiënten



Figuur 1: Grafische weergave van het verband tussen de ratio van de verwijskansen bij ontwikkelde respectievelijk minder ontwikkelde symptomatologie  $\frac{V_1}{V_2}$  en de overall-testkenmerken in de verwezen populatie (Basisinformatie over stratumspecifieke sensitiviteit en specificiteit uit schema 1).

Men kan fouten maken als men de sensitiviteit en de specificiteit in een geselecteerde populatie meet, en deze vervolgens toepast op de prevalentie van een ongeselecteerde populatie ter berekening van voor-

spellende waarden. Aldus zou men, uitgaande van de gegevens in schema 1, de voorspellende waarde van de positieve uitslag in de niet-verwezen populatie onderschatten (71% versus 79%) en die van de negatieve test overschatten (92% versus 87%). Men kan ook, bijvoorbeeld in het kader van een epidemiologische survey, de (door selectie vertekende) sensitiviteit en specificiteit gebruiken om de prevalentie "exact" te berekenen op grond van het aantal gevonden positieve uitslagen (zie de formule in appendix 1). Zo zou men, bij toepassing op de data van schema 1, een prevalentie vinden van slechts 20%, in plaats van de 33% die deze in werkelijkheid bedraagt!

Uit schema 1 wordt tenslotte duidelijk, dat men er goed aan doet, om de testeigenschappen en de voorspellende waarden zowel vóór als ná verwijzing te beschouwen per stratum van ontwikkeling aan de symptomatologie: indien vooral, of, zoals in het voorbeeld, uitsluitend de mate van ontwikkeling van de symptomatologie de verwijskans bepaalt zal er in zo'n stratum weinig of geen verandering van test kenmerken en voorspellende waarden plaatsvinden.

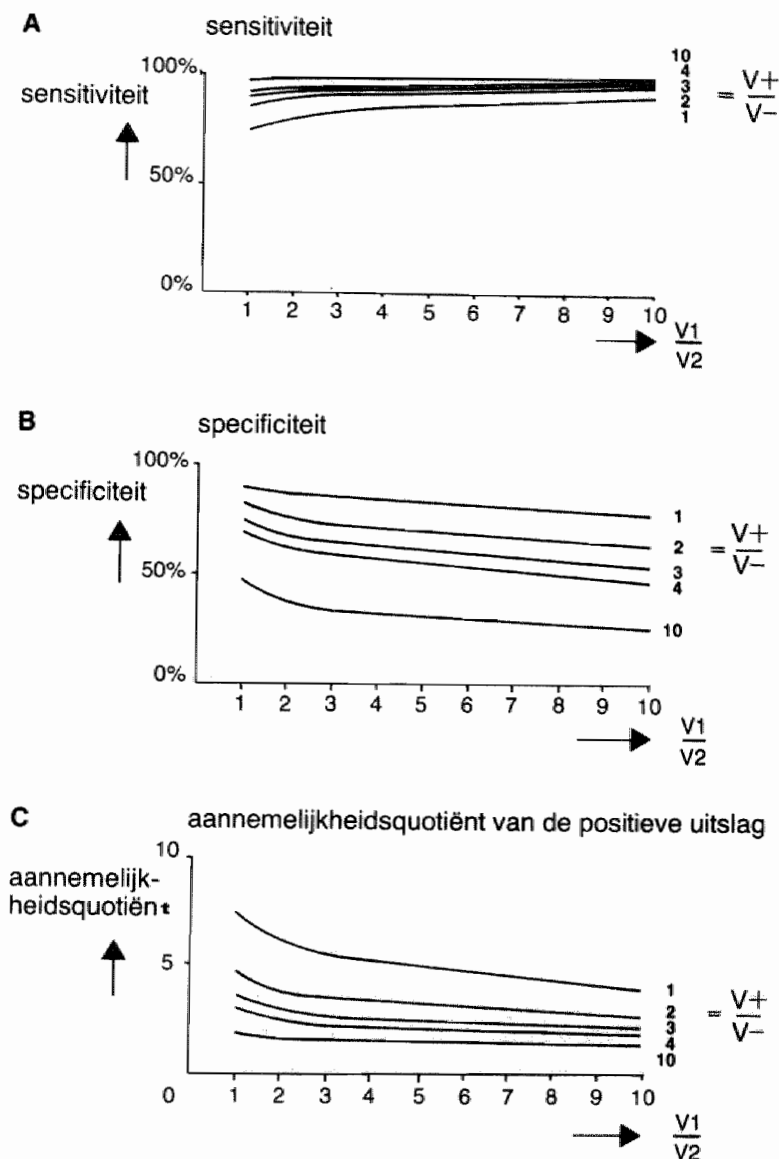
#### Verwijzing naar aanleiding van symptomen en testuitslag.

In de vorige paragraaf is de situatie besproken waarbij alleen verwezen wordt naar aanleiding van de symptomatologie, en niet rechtstreeks in verband met de testuitslag. Het is echter vaak zo dat zowel de symptomatologie als de, reeds door de huisarts vastgestelde, testuitslag een rol spelen bij de selectie door verwijzing. In appendix 2 zijn formules afgeleid voor het verband tussen de overall-sensitiviteit en specificiteit na de verwijzing, als functie van:

- 1) de sensitiviteit en specificiteit vóór de verwijzing, per graad van ontwikkeling van de symptomatologie (twee strata);
- 2) de verhouding tussen de verwijskans bij meer en minder ontwikkelde symptomatologie;
- 3) de verhouding tussen de verwijskans van testpositieven en testnegatieven;
- 4) de verhouding tussen het aantal mensen met meer respectievelijk minder ontwikkelde symptomatologie.

Hierbij is aangenomen dat de invloeden die de symptomatologie en de testuitslag op de verwijskans uitoefenen onafhankelijk van elkaar zijn (indien beide elkaar zouden versterken worden de nu te beschrijven trends nog meer uitgesproken). Men kan nu bestuderen hoe de beschreven

verbanden zich gedragen. We zijn hierbij weer uitgegaan van de situatie in de (ongeselecteerde) populatie uit schema 1.



Figuur 2: Grafische weergave van de invloed op de testkenmerken van de ratio van de verwijskansen bij meer ontwikkelde respectievelijk minder ontwikkelde symptomatologie  $\frac{V_1}{V_2}$  en de ratio van de verwijskansen van test-positieven en test-negatieven  $\frac{V_+}{V_-}$ , indien de verwijskansen in verband met de testuitslag en de symptomatstatus onafhankelijk van elkaar zijn.

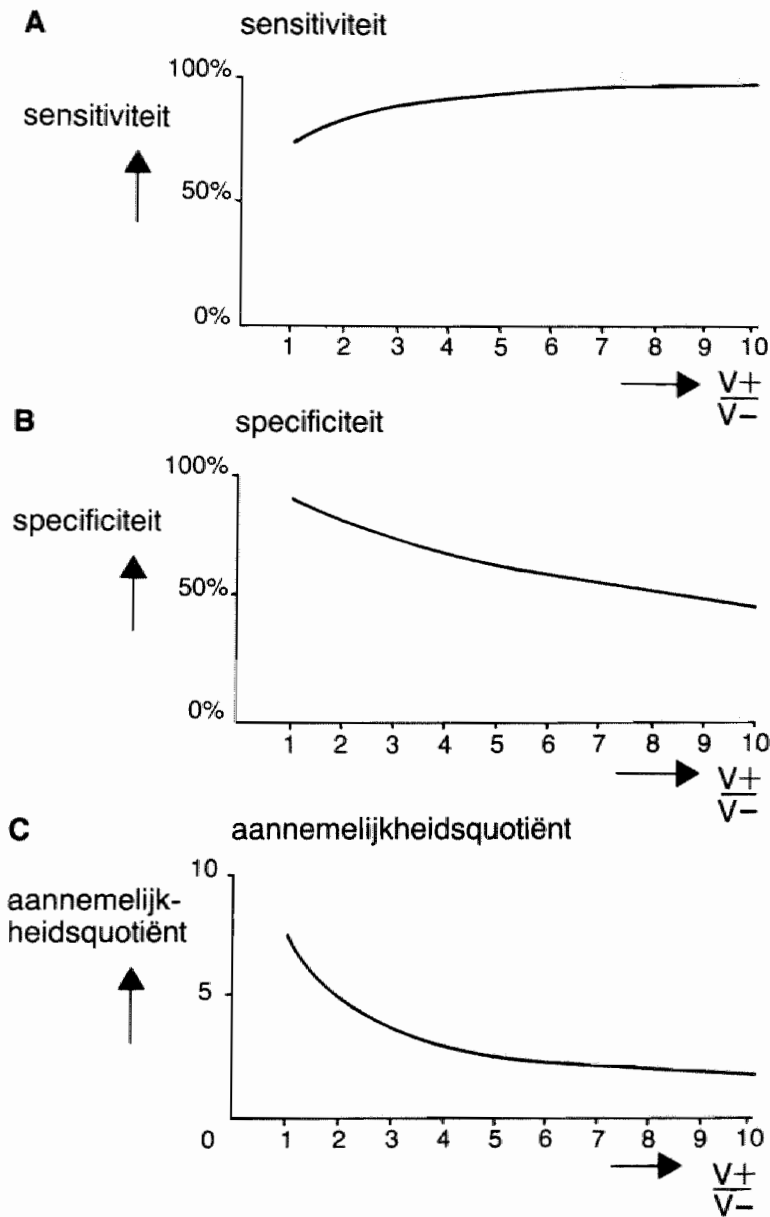
Uit figuur 2 blijkt nu dat, door toevoeging van het element 'selectie naar testuitslag' de trends uit figuur 1 versterkt worden. Bij een toenemende verwijzingskans bij de meer ontwikkelde symptomatologie en positieve testuitslagen zien we de sensitiviteit enigszins toenemen, de specificiteit dramatisch afnemen, en de aannemelijkheidsquotiënten sterk afnemen.

Indien alléén ten aanzien van het al dan niet test-positief zijn wordt geselecteerd, en niet daarnaast nog op de symptomatologie - hetgeen voorkomt als men de testuitslag van overheersend belang vindt ten opzichte van de symptomatologie - dan zullen de voorspellende waarden van de testuitslagen door verwijzing niet veranderen. Wel veranderen de sensitiviteit, de specificiteit en de aannemelijkheidsquotiënten als functie van de verwijzingskans. Voor het geval dat er een positief verband is tussen test-uitslag en ziekte zijn deze veranderingen aangegeven in figuur 3. Het blijkt weer, dat bij een positief verband tussen positieve uitslagen en de verwijzingskans de overall-sensitiviteit toeneemt, de specificiteit afneemt, en de aannemelijkheidsquotiënten afnemen.

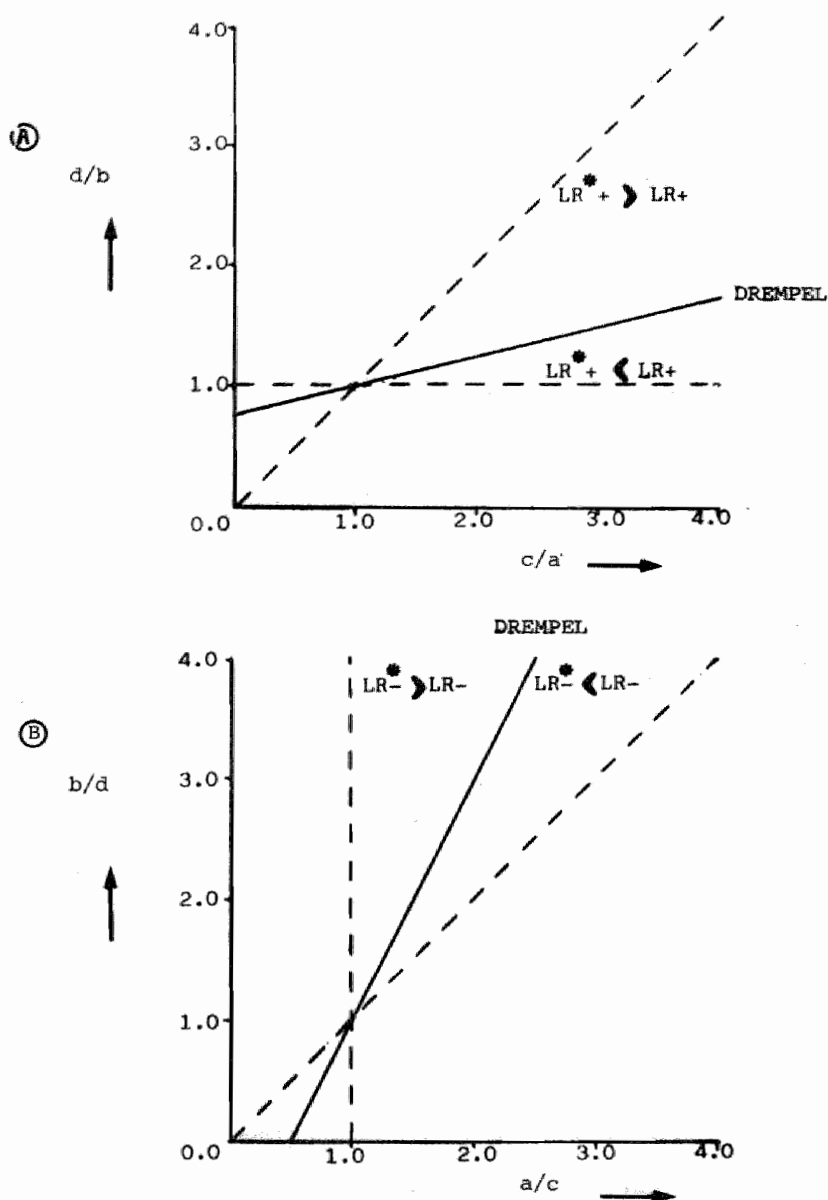
Het onderscheiden van testkenmerken naar de mate van ontwikkeling van de symptomatologie levert in het geval dat de testuitslag de verwijzing beïnvloedt géén 'verwijzings-resistente' testkenmerken per stratum van de symptomatologie op.

#### Een algemene representatie.

In algemene termen gesproken hebben we te maken met mogelijk verschillende overall-verwijzingskansen voor elk van de vier denkbare combinaties van testuitslag- en ziektestatus. Hiervan uitgaande zijn algemene formules af te leiden waarmee is na te gaan in welke richting de aannemelijkheidsquotiënten zullen veranderen bij iedere mogelijke combinatie van die verwijzingskansen. Dat is gebeurd in appendix 3. In figuur 4 is het resultaat grafisch weergegeven. Zo is het bijvoorbeeld te zien dat, bij een positief verband tussen afwijkende uitslagen en de verwijzingskansen en het alleen via de testuitslag beïnvloed worden van de verwijzingskansen (d.w.z.  $a/c = b/d > 1$ ), de aannemelijkheidsquotiënten na verwijzing altijd afgenomen zijn.



Figuur 3: Grafische weergave van de invloed op de testkenmerken van de ratio van de verwijskansen van test-positieven en test-negatieven  $\frac{V+}{V-}$ .



Figuur 4: Relatie tussen overall verwijskansen en de richting van de verandering van de aannemelijkheidsquotiënten  $LR^+*$  (A) en  $LR^-*$  (B), voor een test met in de ongeselecteerde populatie een  $LR^+ = 2,0$  en een  $LR^- = 0,25$ .  
 $a, b, c$  en  $d$  zijn de verwijskansen (c.q. selectiekansen) in een vierveldentabel (zie appendix 3). De gebroken lijnen markeren de mogelijke range van de drempelfunctie voor andere waarden van  $LR^+$  en  $LR^-$ . Bij combinaties van verwijskansen die overeenkomen met punten onder de drempellijnen neemt de LR af. Bij punten boven de lijn neemt de LR toe. De lijn representeert de punten waarbij de LR niet verandert.

Schema 2: Samenhang tussen de mate van ST-depressie op het inspanningselectrocardiogram en het voorkomen van coronaire pathologie (D+) in een (nog) niet, respectievelijk wel door verwijzing geselecteerde populatie. De verwijskans is afhankelijk van de mate van ST-depressie.

(a) Verdeling van patiënten naar mate van ST-depressie vóór en na verwijzing.

ST-depressie	vóór verwijzing		verwijskans	ná verwijzing	
	D+	D-		D+	D-
> 3.0 mm	31	0	100%	31	0
2.5 - 2.9	20	0	75%	15	0
2.0 - 2.4	49	13	55%	27	7
1.5 - 1.9	86	23	35%	30	8
1.0 - 1.4	160	195	20%	32	39
0.5 - 0.9	120	430	10%	12	43
0.0 - 0.4	60	1060	5%	3	53
TOTAAL	526	1721	13%	150	150

b) Uit (a) berekende testkenmerken, vóór en na verwijzing, naar afkappunt.

Testkenmerken 1)		Testkenmerken 1)					Testkenmerken 1)				
Afkappunt		vóór verwijzing					ná verwijzing				
ST-depressie	Se	Sp	LR+	LR-	LRL	Se	Sp	LR+	LR-	LRL	
> 3.0 mm (A)	0.06	1.00		0.94		0.21	1.00		0.79		
2.5 (B)	0.10	1.00		0.90		0.31	1.00		0.69		
2.0 (C)	0.19	0.99	25.2	0.82	12.3	0.49	0.95	9.8	0.54	3.9	
1.5 (D)	0.35	0.98	16.7	0.66	12.2	0.69	0.90	6.7	0.34	3.8	
1.0 (E)	0.66	0.87	5.1	0.39	2.6	0.90	0.64	2.5	0.16	0.8	
0.5 (F)	0.89	0.62	2.3	0.18	0.9	0.98	0.35	1.5	0.06	0.3	
0.0 (G)	1.00	0.00	1.0	-	0.2	1.00	0.00	1.0	-	0.1	

1) Se = sensitiviteit, Sp = specificiteit, LR+ = aannemelijkheidsquotiënt van een positieve uitslag, LR- = aannemelijkheidsquotiënt van een negatieve uitslag, LRL = aannemelijkheidsquotiënt per niveau van de testuitslag.

Tests met meer dan twee uitslag-niveau's.

In de vorige paragrafen zijn we telkens uitgegaan van een op grond van één gefixeerd afkappunt te maken onderscheiding in positieve en negatieve testuitslagen. Bij veel tests zijn er echter meer dan twee uitslagniveau's mogelijk, en in dat geval kan het afkappunt op verschil-



lende niveau's gekozen worden. Daarnaast kan men ook aan ieder uitslagniveau zelf een eigen diagnostische waarde toekennen. Wij zullen naar aanleiding van een uitbreiding van een voorbeeld van Feinstein (8) aandacht besteden aan de mogelijke effecten van verwijspatronen op het onderscheidend vermogen van tests waarbij meer dan twee klassen van uitslagen worden gehanteerd.

In schema 2 is aangegeven hoe de samenhang kan zijn tussen de mate van ST-depressie op het inspanningselectrocardiogram en het voorkomen van angiografisch aangetoonde coronaire pathologie. Deze gegevens zijn vastgesteld door beide onderzoeken te verrichten bij door middel van verwijzing geselecteerde personen. Het ligt voor de hand dat, zeker indien ook huisartsen inspannings-ECG's vervaardigen, er een positief verband is tussen de verwijskans en de mate van ST-depressie. Daarom hebben wij aan de hand van een getallen voorbeeld aangegeven hoe de verdeling van uitslagen onder zieken en niet-zieken zou zijn geweest in de bron populatie die de huisarts consulteerde. Tevens hebben wij aangegeven:

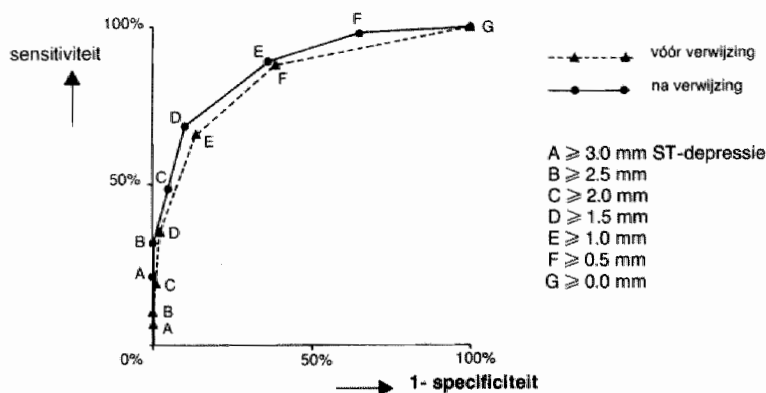
- de sensitiviteit, de specificiteit, en de aannemelijkheidsquotiënten voor de diverse afkappunten in de verwezen en de (nog) niet verwezen populatie;
- de aannemelijkheidsquotiënten voor de diverse uitslag-niveau's afzonderlijk.

Tenslotte is de ROC-curve (die de relatie aangeeft tussen de sensitiviteit en de specificiteit bij diverse afkappunten) voor beide situaties getekend (figuur 5).

Het blijkt, dat indien de verwijzing positief gerelateerd is aan de testuitslag, het patroon van de "dichotome situatie" zich herhaalt: bij elk afkappunt neemt de sensitiviteit toe, en de specificiteit neemt af evenals de aannemelijkheidsquotiënten per uitslag-niveau. De bewering van Sackett et al (16) dat dergelijke aannemelijkheidsquotiënten ongevoelig zouden zijn voor selectie is dus niet juist.

De ROC-curve is minder gevoelig, althans wat haar vorm betreft. Echter de haar bepalende punten schuiven na verwijzing op: om een vergelijkbaar onderscheidend vermogen (c.q. aannemelijkheidsquotient) te krijgen zou men, na verwijzing, een wat hoger afkappunt moeten kiezen,

bijvoorbeeld 1.5 mm in plaats van 1.0 mm. Men moet zich dan wel realiseren dat er meer zieken "gemist" worden. Het minst gevoelig voor selectie, indien deze plaatsvindt vooral of uitsluitend naar aanleiding van het niveau van de uitslag, is weer de voorspellende waarde van een bepaald uitslag niveau: in schema 2 zijn de voorspellende waarden per elevatie-niveau vóór en ná verwijzing gelijk, ondanks de veel hogere overall-prevalentie na verwijzing. Daarbij is wel een voorwaarde, dat na verwijzing de zieken en niet-zieken in de onderzochte groep evenredig (op grond van de samenstelling van de verwezen populatie) vertegenwoordigd zijn.



Figuur 5: ROC-curve (Receiver Operating Characteristic) vóór (▲) en ná (●) verwijzing, uitgaande van de gegevens in schema 2.

### Beschouwing.

Bij verwijzing door de huisarts naar de specialist moet men niet alleen rekening houden met een veranderende (meestal toenemende) prevalentie. Omdat het verwijzgedrag wordt beïnvloed door de symptomatologie, de ernst van de ziekte en door de uitslagen van diagnostisch onderzoek is ook - een soms aanzienlijke - verandering van de validiteit, c.q. het onderscheidend vermogen van diagnostische tests te verwachten.

Het veel beschreven model, dat voorspellende waarden van testuitslagen

in een bepaalde populatie kunnen worden berekend op grond van de (meestal geschatte) prevalentie en de, in andere populaties bepaalde, sensitiviteit en specificiteit, is dan ook te simpel en kan leiden tot vergissingen.

Over verwijspatronen in relatie tot symptomen, aandoeningen en testresultaten is weinig bekend. Hoe groot de besproken invloeden in de praktijk precies zijn voor diverse diagnostica is dan ook op dit moment niet te zeggen. Er is geen empirisch onderzoek gedaan naar de invloed van verwijsgedrag op de validiteit van diagnostische tests en de daarbij werkzame mechanismen. Een probleem bij onderzoek op dit gebied is dat criteria en standaarden voor klachten, symptomen en diagnoses moeilijk zijn te formuleren of op vergelijkbare wijze door verschillende artsen toe te passen. Ingrijpende definitieve diagnostiek in het kader van wetenschappelijk onderzoek is op het niveau van de huisarts in het algemeen niet te effectueren. Niettemin is, ook blijkens het hiervoor besproken, onderzoek naar verwijspatronen zeer relevant.

Wanneer men een indruk heeft van deze verwijspatronen, in relatie tot het klinisch beeld en de resultaten van diagnostische tests, kan men in principe schattingen maken van het onderscheidend vermogen van die tests in (nog) niet door middel van verwijzing geselecteerde populaties. Men doet dan niets anders dan het terugrekenen, uitgaande van de eindtoestand ná selectie, naar de begintoestand vóór selectie. De in appendix 2 gegeven formule geeft een beeld van de variabelen die men voor zo'n schatting in elk geval zou moeten kennen. De tot nu toe in de literatuur beschreven methoden zijn in het algemeen echter nog niet toereikend. Zij berusten op aannames die niet altijd acceptabel zijn, zoals het alleen beïnvloed worden van de verwijskans door testuitslagen en niet ook nog eens door symptomen (17-19). Meer dan een schatting zal een dergelijke berekening bovendien niet kunnen opleveren. Aannames blijven tot op zekere hoogte nodig en bepaalde invloeden worden niet of onvoldoende bekend. Het adagium dat de validiteit van een test eigenlijk het best bepaald kan worden in eenzelfde type populatie als waarin hij toegepast zal worden, blijft dan ook geldig. Zolang dit niet of onvoldoende geschiedt, is het van belang om tenminste een inzicht te hebben in de belangrijkste mechanismen van de selectie, en de richting en de mate waarin zij werkt. Wij hebben laten

zien, dat in de veel voorkomende situatie dat er positieve verbanden bestaan tussen de mate van ontwikkeling van de symptomatologie, de mate van ontwikkeling van de ziekte, het test-positief zijn en de verwijskans, te verwachten is dat:

- de overall-sensitiviteit toeneemt;
- de overall-specificiteit afneemt.

Het aannemelijkheidsquotient van de testuitslag kan dan afnemen.

In feite komt het er op neer, dat het onderscheidend vermogen slechts blijft bestaan voor zover er, bij de verwijzing, niet reeds gebruik is gemaakt van de informatie in de testuitslag.

Indien alléén of vooral de symptomatologie de verwijskans bepaalt, kan men de vertekening door selectie voorkomen of reduceren door de test-kenmerken (sensitiviteit, specificiteit, aannemelijkheidsquotient) te bepalen per graad van ontwikkeling van de symptomatologie. Indien alléén of vooral het niveau van de testuitslag zélf de verwijskans beïnvloedt, is er veel voor te zeggen om de voorspellende waarde per uitslag-niveau als meeste constante kenmerk te beschouwen.

#### LITERATUUR.

1. Sturmans F, Mulder FGH. de betekenis van de prevalentie voor de predictieve waarde van een test. Tijdschr Soc Geneesk 1976; 54: 496-501.
2. Velden HCM van der. Diagnose of prognose: de betekenis van de epidemiologie voor het handelen van de huisarts. Huisarts en Wetenschap 1983; 26: 125-128.
3. Knottnerus JA. Interpretatie van diagnostische gegevens, een onontgonnen terrein in de huisartsgeneeskunde. Huisarts en Wetenschap 1983; 26: 363-368.
4. Baggen J, Leffers P, van Leeuwen YD. Het testdiagram; een visueel middel bij de interpretatie van testuitslagen. Ned Tijdschr Geneesk 1984; 128: 1656-1659.
5. Ransohoff DF, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. New Engl J Med 1978; 299: 926-930.
6. Fletcher RH, Fletcher SW, Wagner EH. Clinical Epidemiology - the essentials. Williams & Wilkins, Baltimore 1982: 51-52.

7. Rümke CL. Kanttekeningen over de gevoeligheid, de specificiteit en de voorspellende waarden van diagnostische tests. Ned Tijdschr Geneesk 1983; 127: 565-561.
8. Feinstein AR. Clinical Epidemiology. The architecture of clinical Research, WB Saunders, Philadelphia 1985.
9. Ottolander GJH den (red.). Interne Geneeskunde 6<sup>e</sup> druk 1980, p. 290.
10. Griner PF, Mayewski RJ, Mushlin AI, Greenland P. Selection and interpretation of diagnostic tests and procedures. Ann Intern Med 1981; 94: 533-596.
11. Green MS. The effect of validation group bias on screeningtests for coronary artery disease. Statistics in Medicine 1985; 4: 53-61.
12. Meischke-de Jongh ML, Blonk DI, Gan-Siauw IN, den Hoed-Sijtsema S, ter Laag M. Borstklachten en hun betekenis voor de diagnostiek van het mammacarcinoom. Ned Tijdschr Geneesk 1983; 127: 2361-2367.
13. Meischke-de Jongh ML, Gan-Siauw IN, den Hoed-Sijtsema S, Blonk DI, Fitskie-de Ronde EHCM, ter Laag M. De waarde van echografie bij het mamma-onderzoek. Ned Tijdschr Geneesk 1985; 129: 1339-1347.
14. Dronkers DJ. Mammografisch occult carcinoom. Ned Tijdschr Geneesk 1985; 129: 1632-1635.
15. Kottnerus A, Sturmans F. Over vertekening van test-uitslagen door reeds aanwezige informatie. The Practitioner 1985; 3: 287-295.
16. Department of Clinical Epidemiology & Biostatistics, Manual for a workshop on how to teach the critical appraisal of clinical evidence. Hamilton Ontario, 1984.
17. Dawid AP. Properties of diagnostic data distributions. Biometrics 1983; 39: 207-216.
18. Begg CB, Greenes RA. Assessment of diagnostic tests when disease verification is subject to selection bias. Biometrics 1983; 39: 207-216.
19. Gray R, Begg CB, Greenes RA. Construction of Receiver Operating Characteristic curves when disease verification is subject to selection bias. Med Decis Making 1984; 4: 151-163.

# Appendix 1.

Berekening van de prevalentie uit het percentage afwijkende uitslagen in een onderzochte populatie.

We gaan uit van de situatie, zoals weergegeven in de volgende viervelden tabel:

	ziek	niet-ziek	
positieve testuitslag	sensitiviteit x prevalentie	(1-specificiteit) x (1-prevalentie)	$P(T^+)$
negatieve testuitslag	(1-sensitiviteit) x prevalentie	specificiteit x (1-prevalentie)	$P(T^-)$
	prevalentie	1-prevalentie	1.0

Nu geldt:

$$P(T^+) = (\text{sensitiviteit} \times \text{prevalentie}) + (1 - \text{specificiteit}) \times (1 - \text{prevalentie})$$

hieruit volgt:

$$\text{Prevalentie} = \frac{P(T^+) + \text{specificiteit} - 1}{\text{sensitiviteit} + \text{specificiteit} - 1}$$

## Appendix 2:

De overall-sensitiviteit, -specificiteit, en -aannemelijkheidsquotiënten kunnen worden uitgedrukt als functie van

- de testkenmerken (sensitiviteit en specificiteit) per stratum van ontwikkeling van de symptomatologie S.
- de verwijskans van test-positieven en test-negatieven.
- de verwijskans bij een meer ontwikkelde en bij een minder ontwikkelde symptomatologie.
- de verhouding van het aantal personen met meer, respectievelijk minder ontwikkelde symptomatologie in de (nog) niet-verwezen populatie.

We gaan uit van de volgende situatie in de oorspronkelijke populatie:

Graad van ontwikkeling van de <u>symptomatologie</u>	Ziekte status	
	<u>Ziek</u>	<u>Niet-ziek</u>
$S_1$	$Se(nsitiviteit)_1$	$Sp(ecificiteit)_1$
	$N_1$	$N_3$
	$V_{1+}$	$V_{3+}$
	$V_{1-}$	$V_{3-}$
$S_2$	$Se(nsitiviteit)_2$	$Sp(ecificiteit)_2$
	$N_2$	$N_4$
	$V_{2+}$	$V_{4+}$
	$V_{2-}$	$V_{4-}$

waarin:

$sensitiviteit_1$  = sensitiviteit bij symptomatologie-status  $S_1$  ( $Se_1$ )

$specificiteit_1$  = specificiteit bij symptomatologie-status  $S_1$  ( $Sp_1$ )

$N_1$  = het aantal personen met een bepaalde symptomatologie- en ziekte-status

$V_{i+}$  = de verwijskans van testpositieven met symptomatologie- en ziekte-status i

$V_{i-}$  = de verwijskans van testnegatieven met een bepaalde symptomatologie- en ziekte-status

Nu geldt:

.A de sensitiviteit in de door verwijzing geselecteerde populatie, sensitiviteit'

$$= \frac{(Se_1 \cdot N_1 \cdot V_{1+}) + (Se_2 \cdot N_2 \cdot V_{2+})}{(Se_1 \cdot N_1 \cdot V_{1+}) + (Se_2 \cdot N_2 \cdot V_{2+}) + [(1-Se_1) \cdot N_1 \cdot V_{1-}] + [(1-Se_2) \cdot N_2 \cdot V_{2-}]} \quad (1)$$

$$= \frac{1}{1 + \frac{(1-Se_1) \cdot N_1 \cdot V_{1-} + (1-Se_2) \cdot N_2 \cdot V_{2-}}{Se_1 \cdot N_1 \cdot V_{1+} + Se_2 \cdot N_2 \cdot V_{2+}}} \quad (2)$$

Indien de verwijzingskansen in verband met de testuitslag en de symptoomstatus onafhankelijk van elkaar zijn, dan geldt:

$$\frac{V_{1+}}{V_{1-}} = \frac{V_{2+}}{V_{2-}} = C_1 \text{ en: } \frac{V_{1+}}{V_{2+}} = \frac{V_{1-}}{V_{2-}} = C_2$$

de formule (2) is nu te herleiden tot:

$$\text{sensitiviteit}' = \frac{1}{1 + \frac{1}{C_1} \cdot \frac{(1-Se_1) + (1-Se_2)/C_2 \cdot q}{Se_1 + Se_2/C_2 \cdot q}} \quad (3)$$

waarbij  $C_1$  = de ratio van de verwijzingskansen van test-positieve en test-negatieve zieken

$C_2$  = de ratio van de verwijzingskansen van zieken met een meer ontwikkelde resp. een minder ontwikkelde symptomatologie

en  $q$  = de ratio van het oorspronkelijke aantal zieken met symptoom status  $S_1$  en het aantal zieken met symptoomstatus  $S_2$ .

.B Op analoge wijze is af te leiden, dat de specificiteit in de door verwijzing geselecteerde populatie, specificiteit'

$$= \frac{1}{1 + \frac{(1-Sp_1)N_3 \cdot V_{3+} + (1-Sp_2)N_4 \cdot V_{4+}}{(Sp_1 \cdot N_3 \cdot V_{3-}) + (Sp_2 \cdot N_4 \cdot V_{4-})}} \quad (4)$$

Indien de verwijzing in verband met de testuitslag en de symptoomstatus onafhankelijk van elkaar zijn, dan geldt weer:



$$\frac{V3+}{V3-} = \frac{V4+}{V4-} = C3 \quad \text{en} \quad \frac{V3+}{V4+} = \frac{V3-}{V4-} = C4$$

formule (4) vereenvoudigt dan tot:

$$= \frac{1}{1 + C3 \cdot \frac{(1-Sp_1) + (1-Sp_2)/C4 \cdot r}{Sp_1 + Sp_2/C4 \cdot r}} \quad (5)$$

waarbij: C3 = de ratio van de verwijskans van test-positieve en test-negatieve niet-zieken

C4 = de ratio van de verwijskans van niet-zieken met een meer resp. minder ontwikkelde symptomatologie.

en r = de ratio van het oorspronkelijke aantal niet-zieken met symptomstatus  $S_1$  en het aantal niet-zieken met symptomstatus  $S_2$ .

Indien de werkelijke ziektestatus vóór verwijzing nog niet bekend is, is het aannemelijk dat:

$$C1 = C3 = \frac{V+}{V-} \quad \text{en} \quad C2 = C4 = \frac{V1}{V2}$$

.C Het aannemelijkheidsquotiënt van de positieve uitslag na verwijzing is nu:

$$\frac{\text{Sensitiviteit}'}{1 - \text{Specificiteit}'}$$

en van een negatieve uitslag:

$$\frac{1 - \text{Sensitiviteit}'}{\text{Specificiteit}'}$$

### Appendix 3

Stel, voor de testpositieve zieken, de testpositieve niet-zieken, de testnegatieve zieken en de testnegatieve niet-zieken gelden respectievelijk de verwijskansen a, b, c, en d.

Als we de sensitiviteit en de specificiteit in de ongeselecteerde populatie Se respectievelijk Sp noemen, geldt:

$$LR+^* = LR+ \cdot \frac{(1-Sp) + Sp.d/b}{(Se + (1-Se).c/a)} \quad (1)$$

en:

$$LR-^* = LR- \cdot \frac{(1-Sp) \cdot b/d + Sp}{Se.a/c + (1-Se)} \quad (2)$$

waarbij LR+ en LR- de aannemelijkheidsquotiënten, respectievelijk voor de positieve en de negatieve uitslag, in de ongeselecteerde populatie representeren, en

LR+\* en LR-\* de overeenkomstige quotiënten in de door middel van verwijzing geselecteerde populatie.

Als we als maat voor de vertekening van het aannemelijkheidsquotiënt ten gevolge van verwijzing nemen de verhouding tussen het aannemelijkheidsquotiënt in de geselecteerde en de ongeselecteerde populatie, dan geldt:

$$\frac{LR+^*}{LR+} = \frac{(1-Sp) + Sp.d/b}{Se + (1-Se).c/a} \quad (3)$$

Om, in algemene termen, de richting van de vertekening te beschrijven, kunnen wij, als functie van de verwijskansen, de drempel (threshold) functie afleiden door  $LR+^* = LR+$  te stellen. Na het vervangen van de linkerterm van (3) door 1.0 en onder uitvoering van enkele eenvoudige algebraïsche substituties, blijkt deze functie equivalent aan:

$$(\hat{d}/\hat{b} - 1) = LR- \cdot (\hat{c}/\hat{a} - 1) \quad (4)$$

waarbij  $\hat{d}$ ,  $\hat{b}$ ,  $\hat{c}$  en  $\hat{a}$  die realisaties van de verwijskansen zijn waarvoor  $LR+^* = LR+$ , c.q. waarbij er geen vertekening optreedt.

Op analoge wijze blijkt de drempelfunctie van LR-\* en LR- te zijn:

$$(b/\bar{d} - 1) = LR+ \cdot (\bar{a}/\bar{c} - 1) \quad (5)$$

waarbij er voor de realisaties  $\bar{b}$ ,  $\bar{d}$ ,  $\bar{a}$  en  $\bar{c}$  geen vertekening van LR-\* ten opzichte van LR- optreedt.

(4) en (5) drukken de drempels uit als lineaire relaties tussen respectievelijk  $(\hat{d}/\hat{b} - 1)$  en  $(\hat{c}/\hat{a} - 1)$  en tussen  $(b/\bar{d} - 1)$  en  $(\bar{a}/\bar{c} - 1)$ , gecentreerd op punt (1, 1) met als lineaire regressiecoëfficiënten respectievelijk LR- en LR+.

Met behulp van figuur 4 kunnen we de richting van de vertekening nu gemakkelijk evalueren, als we uitgaande van de originele aanneemelijkheidsquotiënten LR- en LR+ de grafiek door (1, 1) tekenen en hierin de (waarschijnlijke) schattingen van de verwijskansen voor elk van de categorieën inbrengen. In geval van een werkelijke of geschatte combinatie van a, b, c en d welke een punt boven de rechte representeert, is het aanneemelijkheidsquotiënt na verwijzing groter geworden. Voor punten onder de rechte geldt dat het aanneemelijkheidsquotiënt na verwijzing kleiner is geworden.

DEEL V

ENKELE OBSERVATIONELE ONDERZOEKINGEN NAAR HET VERBAND  
TUSSEN UITSLAGEN VAN DIAGNOSTISCHE TESTS EN PROGNOSE.

HEMOGLOBINE EN ZWANGERSCHAPSUITKOMST.

J.A. Knottnerus 1) - L. Delgado 2) - P.G. Knipschild 2) - G.G.M. Essed 3) - F. Smits 3)

Samenvatting.

Beschreven wordt een onderzoek naar het verband tussen de waarden van haematologische parameters tijdens de zwangerschap, het geboortegewicht en de zwangerschapsduur.

Bij 494 zwangeren werden rond de 30e zwangerschapsweek de haematologische parameters bepaald, en werd door middel van een vragenlijst informatie ingewonnen over rookgedrag, alcoholgebruik, en gebruik van staaltabletten. Gevonden werd een verband tussen enerzijds een hemoglobinegehalte (Hb) van 8.0 mmol/l en hoger en een hematocriet (Ht) van 39% en hoger, en anderzijds de kans op een laag geboortegewicht en prematuritas. Dit verband bleef aanwezig indien gecorrigeerd werd voor roken, alcoholgebruik, pariteit en het gebruik van staaltabletten. Voor Hb- en Ht-waarden gemeten vóór de 16e zwangerschapsweek kon geen duidelijk verband worden gevonden. De mate van hemodilutie hangt sterk samen met de uitgangswaarde. De resultaten zijn in overeenstemming met de hypothese dat een te geringe mate van hemodilutie, cq. het niet bereiken van de optimale range, een ongunstige invloed heeft op de placenta-doorbloeding.

- 1) Capaciteitsgroep Epidemiologie en Capaciteitsgroep Huisartsgeneeskunde
  - 2) Capaciteitsgroep Epidemiologie
  - 3) Capaciteitsgroep Verloskunde en Gynaecologie
- Rijksuniversiteit Limburg

## Inleiding.

In hoofdstuk 8 is verslag gedaan van een onderzoek naar het verband tussen hemoglobinegehalte (Hb) van zwangeren gemeten rond de dertigste zwangerschapsweek, en subjectieve klachten. Het bleek dat bij hemoglobine waarden van 7.0 mmol/l of lager bepaalde klachten vaker voorkwamen dan bij hogere Hb-waarden. De hiervan af te leiden (algemene) ondergrens voor het Hb-gehalte tijdens de hemodilutiefase van de zwangerschap is hoger dan die welke op grond van een optimaal verloop van de zwangerschap wordt aanbevolen (1, 2). Daarnaast is er de laatste jaren toenemende aandacht gekomen voor de mogelijke nadelen van een (te) hoog Hb-gehalte (1 - 6), casu quo een te geringe mate van plasma volume toename en hemodilutie (7 - 10) tijdens de zwangerschap. De zone waarbinnen het Hb-gehalte optimaal zou zijn, wordt daarmee tot een betrekkelijk smal gebied beperkt. Inmiddels zijn er nog kanteekeningen te maken bij het onderzoek naar de relatie hoog Hb - zwangerschapsuitkomst. De meeste studies betroffen erg kleine en geselecteerde groepen (5 - 9) en slechts enkele onderzoekers hielden rekening met de mogelijke invloed van versturende variabelen, zoals roken (1, 4) of bloeddruk (4). Bij het grote onderzoek van Garn (1), die een U-vormig verband vond tussen Hb en hematocriet (Ht) enerzijds en de ongunstige uitkomst anderzijds, is bovendien enige selectiebias niet uitgesloten. Men hield geen rekening met het aantal weken amenorroe waarbij het Hb bepaald werd. Dit kan bijvoorbeeld betekenen dat men bij een partus immaturus een Hb-waarde uit de vroege fase van zwangerschap in de analyse betrokken heeft. Omdat het Hb dan nog weinig gedaald hoeft te zijn, kan er een schijnassociatie optreden tussen een hoog Hb en een ongunstige uitkomst.

In het volgende doen wij verslag van een onderzoek naar de invloed van hoge Hb- en Ht(hematocriet)-waarden, vastgesteld bij de zwangerschapscontrole rond dertig weken, op de zwangerschapsuitkomst, rekening houdend met de mogelijke invloeden van roken, alcoholgebruik, bloeddruk, pariteit en het gebruik van staaltabletten.

## Methode.

In de periode van 1 juli 1985 tot en met eind april 1986 werd een vragenlijst voorgelegd aan de zwangeren die de polikliniek verloskunde

van het Academisch Ziekenhuis Maastricht bezochten in verband met de eerste zwangerschapscontrole na de dertigste week. Gevraagd werd onder meer naar roken en alcoholgebruik gedurende respectievelijk de eerste drie maanden, de tweede drie maanden en de voorafgaande weken, en naar het gebruik van staaltabletten. De vragenlijst werd vóór het verlaten van de polikliniek ingeleverd. Direct aansluitend werd bloed afgenomen door middel van venapunctie. Het bloed werd opgevangen in EDTA-buizen, en binnen enkele uren werden de hematologische parameters bepaald met een SYSMEX Microcell-counter (hematologisch laboratorium Academisch Ziekenhuis Maastricht). Na afloop van de zwangerschap werden, op een standaardformulier, uit de status gegevens genoteerd over het verloop van de zwangerschap (onder andere bloeddruk, het wel of niet aanwezig zijn van een primaire of secundaire medische indicatie, de pariteit) en de uitkomst (uiteindelijke zwangerschapsduur en geboortegewicht). De vragenlijst werd ook voorgelegd aan de zwangeren die begeleid werden in twee perifere verloskundige praktijken, bij de eerste controle ná 30 weken amenorroe in de periode van 1 oktober 1985 tot 1 februari 1986. Direct aansluitend werd venapunctie verricht in de verloskundige praktijk of op één van de perifere prikpunten van de Prikdienst van het Academisch Ziekenhuis Maastricht. Ook van deze groep zwangeren werden post partum gegevens verzameld over het verdere verloop van de zwangerschap. De geboortegewichten op de verloskamer werden vastgesteld met een speciaal geijkte electronische weegschaal (Berkel) terwijl de zelfstandig gevestigde verloskundigen het geboortegewicht bepaalden met, wederom speciaal geijkte, veer-weegschalen van het type Unster. In beide gevallen werden dezelfde ijkgewichten gebruikt. De gegevens werden bewerkt op de VAX-computer van de Rijksuniversiteit Limburg en geanalyseerd met het BMDP-programma. In de analyse zijn betrokken de bloeddrukwaarden bij de hoogst gemeten diastolische bloeddruk. Deze werd in 90% van de gevallen in het derde trimester vastgesteld. De statistische analyse vond als volgt plaats. Eerst werden de verbanden tussen de hematologische parameters, geboortegewicht en zwangerschapsduur bestudeerd met behulp van kruistabellen, puntendiagrammen en univariate lineaire regressie analyse. Ook werden onderzocht de verbanden tussen roken, alcoholgebruik, pariteit en diastolische bloeddruk enerzijds en hematologische parameters, geboortegewicht en zwangerschapsduur anderzijds. Op grond van de

bevindingen werden multipele regressie modellen opgesteld, met een laag geboortegewicht en/of vroeggeboorte als onafhankelijke variabelen. Omdat vooral bij hoge waarden van Hb en Ht een verband met de uitkomstparameters werd gevonden en op grond van de literatuur, werd een kwadratisch model ontworpen dat gebaseerd is op het in principe bestaan van een U-vormig verband tussen de hematologische parameters en de uitkomstparameters. Hierbij is het Hb, respectievelijk het Ht, getransformeerd door er de waarde van af te trekken waarbij op grond van de data het laagste punt van de "U" wordt verondersteld. Voor het Hb is dit de waarde 7.9 mmol/l en voor de hematocriet 37%. Als men in het model een kwadratische term voor het Hb, respectievelijk de Ht toevoegt, ontstaat een U-vormige functie.

Voorts werd nagegaan in hoeverre het al dan niet gebruikt hebben van staaltabletten van invloed was en of er verschillen waren tussen de diverse bestudeerde subpopulaties. Tenslotte werd de samenhang tussen de hematologische uitslagen bij de eerste en de dertig-weeken controle bestudeerd, in relatie tot de zwangerschapsuitkomst.

### Resultaten.

Van de zwangeren die via de polikliniek verloskunde benaderd werden, nam 92% deel aan het onderzoek. Redenen voor niet-deelname waren: weigering (4%), taalbarrière (2%), het niet op de betreffende controle verschijnen (2%). Van de zwangeren die via de zelfstandig gevestigde verloskundigen benaderd werden nam 63% deel; 12% weigerde, 10% vond de afstand naar de prikdienst te groot, 8% werd bij de controle gemist, en bij 7% waren er diverse andere redenen.

Voor de analyse in het kader van dit proefschrift waren beschikbaar de gegevens van 494 zwangeren (90 "perifeer" begeleid, en 404 poliklinisch). De leeftijdsrange was 18 tot 39 jaar (P50=27), en 50% was nullipara. Bij alle deelnemende vrouwen is het Hb bepaald (gemiddeld 7.3 mmol/l, SD: 0.6), bij 96% de hematocriet (gemiddeld 35%, SD: 3) en bij 91% het MCV (gemiddeld 91, SD: 5).

Omdat de non-respons bij zwangeren die via de vrijgevestigde verloskundigen benaderd werden, vrij hoog was (37%), gingen wij na of er sprake was van selectieve uitval ten aanzien van zwangerschapsduur en/of geboortegewicht. Daarvoor konden wij geen aanwijzingen vinden. Het gemiddelde geboortegewicht in de non-respons groep was 3328 gram



(SD: 564 gram) versus 3316 gram (SD: 494 gram) in de deelnemende groep. Voor de zwangerschapsduur waren de overeenkomstige cijfers 39 weken en 4 dagen (SD: 1 week en 5 dagen) respectievelijk 39 weken en 4 dagen (SD: 1 week en 4 dagen).

Tabel 1. Hemoglobinegehalte tussen 30 en 32 weken amenorroeduur, zwangerschapsuitkomst en bloeddruk.

Hemoglobine (mmol/l) <sup>1)</sup>	<6.4		6.5-6.9		7.0-7.4		7.5-7.9		>8.0	
	n	(%)	n	(%)	n	(%)	n	(%)	n	(%)
geboortegewicht										
< 2500 gram	1	(3)	6	(6)	6	(4)	3	(2)	8	(13) <sup>3)</sup>
< 3000 gram	7	(24)	29	(27)	36	(22)	29	(21)	21	(34) <sup>4)</sup>
vroeggeboorte (<37 weken)	1	(3)	3	(3)	9	(6)	1	(3)	6	(10) <sup>4)</sup>
hoge bloeddruk <sup>2)</sup>	3	(10)	5	(5)	14	(9)	22	(16)	13	(21) <sup>4)</sup>
totaal	29	(100)	107	(100)	161	(100)	135	(100)	62	(100)

1) 1 mmol/l = 1.61 g/dl.

2) diastolische druk > 90 mmHg en/of systolische druk > 140 mmHg bij gelegenheid van de hoogst gemeten diastolische bloeddruk tijdens de prenatale zorg.

3)  $p < 0.01$  (de hoogste categorie vergeleken met de rest).

4)  $p < 0.05$  (de hoogste categorie vergeleken met de rest).

(Bij de statistische analyse werd de chikwadraat toets gebruikt, of de Fisher exact test als het verwachte aantal patiënten in één van de (vier) cellen 5 of lager bedroeg).

In de tabellen 1 en 2 worden de relaties tussen Hb en Ht enerzijds en geboortegewicht, zwangerschapsduur en bloeddruk anderzijds gepresenteerd. Het is duidelijk dat hoge Hb- en Ht-waarden ( $> 8.0$  mmol/l respectievelijk  $> 39\%$ ) samenhangen met een laag geboortegewicht en prematuritas. Een hogere bloeddruk vinden we al boven een Hb van 7.5 mmol/l en een Ht van 35%.

Voor lage Hb- en Ht-waarden vonden wij geen duidelijke verschillen, mogelijk omdat erg lage waarden zeldzaam waren. Voor het MCV (Mean Corpuscular Volume) vonden wij geen relevante invloed op de zwangerschapsuitkomst. Blijkbaar is het erythrocytengetal sterker van invloed dan het celvolume.

Tabel 2. Hematocriet tussen 30 en 32 weken amenorroeduur, zwangerschapsuitkomst en bloeddruk.

Hematocriet (%)	<30		31-34		35-38		>39	
	n	(%)	n	(%)	n	(%)	n	(%)
geboortegewicht								
< 2500 gram	2	(5)	10	(5)	4	(2)	7	(14)
< 3000 gram	10	(24)	41	(22)	51	(26)	14	(27)
vroeggeboorte (<37 weken)	1	(2)	9	(5)	3	(2)	5	(10)
hoge bloeddruk <sup>1)</sup>	4	(10)	15	(8)	27	(14)	29	(18)
totaal	42	(100)	186	(100)	196	(100)	51	(100)

1) diastolische druk > 90 mmHg en/of systolische druk > 140 mmHg bij gelegenheid van de hoogst gemeten diastolische bloeddruk tijdens de prenatale zorg.

2)  $p < 0.01$  (de hoogste categorie vergeleken met de rest).

3)  $p < 0.05$  (de hoogste categorie vergeleken met de rest).

(Bij de statistische analyse werd de chikwadraat toets gebruikt, of de Fisher exact test als het verwachte aantal patiënten in één van de (vier) cellen 5 of lager bedroeg).

Roken had een sterke negatieve invloed op het geboortegewicht, waarbij in tweeërlei zin sprake was van een dosis-respons relatie: naarmate men later in de zwangerschap nog rookt, en naarmate men méér sigaretten rookt is het effect sterker. Van wie rond de dertigste week nog rookte, bleek het kind per gerookte sigaret per dag een verlies van 24 gram geboortegewicht ten opzichte van niet (inclusief ex-)rokers te hebben. Bij rookers werd ook vaker vroeggeboorte gevonden: bij 6% van hen duurde de zwangerschap korter dan 37 weken versus 3% bij niet-rookers ( $p=0.08$ ). Het verband tussen Hb en Ht en zwangerschapsuitkomst kon niet door roken verklaard worden: onafhankelijk van de hoogte van het Hb en de Ht rookte 35% van de zwangeren rond de dertigste week. Alcoholgebruik vertoonde evenmin verband met de hematologische parameters en wij konden ook geen systematische invloed ervan op de uitkomstparameters vaststellen.

Opmerkelijk is, dat er geen duidelijke verbanden konden worden gevonden tussen de hoogst gemeten diastolische bloeddruk enerzijds, en het geboortegewicht en de zwangerschapsduur anderzijds. Een mogelijke verklaring zou zijn, dat bij de prenatale controle direct wordt inge-

speeld op verhoogde bloeddruk, zodat ongunstige effecten later moeilijk te meten zijn.

Bij hogere pariteit werd vaker een laag Hb aangetroffen: van de nulliparae, primiparae en de multiparae had respectievelijk 4% (10/238), 6% (11/181) en 11% (8/75) een Hb lager dan 6.5 mmol/l. Het gemiddelde geboortegewicht nam in deze groepen toe: respectievelijk 3249 gram (SD:497), 3352 gram (SD: 480), 3440 gram (SD: 495), zonder dat er verschillen in zwangerschapsduur werden waargenomen.

Vrouwen die staaltabletten gebruikten of hadden gebruikt, hadden rond de dertigste zwangerschapsweek een iets lager Hb dan de anderen (7.2 mmol/l versus 7.3 mmol/l), doch de geboortegewichten en zwangerschapsduren verschilden niet duidelijk. Overigens bleek bij de vrouwen die staal gebruikten, het gemiddelde Hb bij de éérste zwangerschapscontrole relatief nog lager te zijn geweest (7.8 versus 8.2 mmol/l,  $p < 0.001$ ). Blijkbaar is er in het verloop van de zwangerschap niet alleen routinematig, maar zeker ook op indicatie, ijzer verstrekt.

In tabel 3 worden de resultaten van de multiële regressie-analyse weergegeven. Afhankelijke variabele is de kans op een geboortegewicht lager dan 2500 g en/of een zwangerschapsduur van minder dan 37 weken. Het blijkt dat roken en hematocriet ook na correctie voor de invloed van de andere variabelen, een duidelijke invloed hebben. De invloed van het Hb bereikt de statistische significantie drempel ( $p < 0.05$ ) niet. Dit is wel het geval als men de analyse beperkt tot de zwangeren die géén staal gebruikt hebben. De reden hiervoor kan zijn dat onder vrouwen die staal gebruikt hebben, minder vaak hoge Hb-waarden worden aangetroffen.

Het verband tussen de zwangerschapsuitkomst en de hematologische parameters verandert niet indien men de analyse apart verricht voor vrouwen die primair door verloskundigen begeleid zijn en voor vrouwen die vanaf het begin door de gynaecoloog zijn begeleid.

De kans op het krijgen van een secundaire medische indicatie voor de bevalling stijgt met het hoger worden van het Hb en de Ht. Dit is vooral duidelijk in de hoogste Hb- en Ht-klassen (tabel 4).

Tabel 3. Multipele regressie, met als afhankelijke variabele de kans op een geboortegewicht lager dan 2500 gram en/of een zwangerschapsduur van minder dan 37 weken (additieve modellen).

variabele	eenheid	(1)			(2)		
		coëfficiënt	SE	P	coëfficiënt	SE	P
(Ht-37) <sup>2</sup>	%	0.003	0.001	0.001 *			
(Ht-37)	%	0.014	0.006	0.008			
(Hb-7.9) <sup>2</sup>	mmol/l				0.041	0.025	0.098
(Hb-7.9)	mmol/l				0.058	0.036	0.106
ROKEN	sigaretten per dag	0.007	0.002	0.001 *	0.007	0.002	0.002 *
HOOGSTE DIAS-TOLISCHE DRUK	mmHg	0.002	0.001	0.198	0.001	0.001	0.262
PARITEIT	nulli = 0 > primi = 1	-0.022	0.032	0.500	-0.023	0.033	0.484
LIJZER-MEDICATIE	nee = 0 ja = 1	-0.009	0.023	0.685	-0.012	0.024	0.633
INTERCEPT		-0.071			-0.348		

Voorbeeld voor de interpretatie: zwangere met hematocrieten van 31% en 43% met overigens vergelijkbare waarden voor de overige variabelen hebben ten opzichte van een zwangere met een hematocriet van 37%

respectievelijk een extra risico van: (31-37)<sup>2</sup> x 0.003 + (31-37) x 0.014 = 2.4% en

(43-37)<sup>2</sup> x 0.003 + (43-37) x 0.014 = 19.2%

(uit vergelijking (1)).

\*: statistisch significant (P < 0.01)

Tabel 4. Het Hb-gehalte en de Ht rond de dertigste zwangerschapsweek in relatie tot een primaire respectievelijk secundaire medische indicatie. Er is horizontaal gepercenteerd.

	geen medische indicatie	primaire medische indicatie	secundaire medische * indicatie	totaal n
<u>Hb (mmol/l)</u>				
-6.4	35%	52%	14%	29
6.5-6.9	27%	54%	19%	107
7.0-7.4	34%	42%	24%	160
7.5-7.9	30%	44%	26%	136
8.0-8.4	33%	44%	24%	55
8.5+	0%	43%	57%	7
totaal	31%	46%	23%	494
<u>Ht (%)</u>				
-30	29%	55%	17%	42
31-32	32%	51%	17%	75
33-34	25%	53%	22%	111
35-36	36%	39%	25%	124
37-38	31%	43%	26%	72
39-40	30%	43%	28%	40
41+	18%	46%	36%	11
totaal	31%	47%	23%	475

\* De lineaire regressiecoëfficiënten van het Hb en de Ht voor de kans op een secundaire medische indicatie als afhankelijke variabele zijn respectievelijk 0.09 (SE: 0.033) en 0.012 (SE: 0.06).

Voor de vrouwen die aanvankelijk door perifere verloskundigen begeleid waren, ontbrak een gestandaardiseerde meting van de aanvangs Hb- en Ht-waarden. Daarom bestudeerden wij alleen voor de geheel poliklinisch begeleide zwangeren (door de poliklinisch werkzame verloskundigen en/of de gynaecologen), of het negatieve verband tussen de hoogte van de hematologische parameters en het geboortegewicht respectievelijk de zwangerschapsduur ook al aan te tonen was wat betreft het Hb en Ht, gemeten bij de éérste zwangerschapscontrole (bij een mediane amenorroeduur van 9 weken; vrouwen die na de zestiende week voor het eerst op controle kwamen werden van deze analyse uitgesloten). Dergelijke verbanden werden voor de hematocriet niet gevonden. Voor het Hb werden evenmin statistisch significante verbanden gevonden, maar was er wel een zwakke U-vormige trend te bespeuren wat betreft de kans op een geboortegewicht onder de 2500 gram (tabel 5). Hoewel de aantallen te klein zijn voor een definitief oordeel, lijkt de conclusie voorlopig gewettigd, dat de eventuele invloed van het Hb en de Ht bij de eerste zwangerschapscontrole in elk geval minder uitgesproken is dan de invloed van deze parameters bij 30 weken.

Tabel 5. Hemoglobinegehalte en hematocriet vóór de 16e zwangerschapsweek, geboortegewicht en zwangerschapsduur.

Hemoglobine (mmol/l) <u>n = 326</u>	<7.4	7.5-7.9	8.0-8.4	8.5-8.9	≥9.0
n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
geboortegewicht < 2500 gram	3 (7)	4 (5)	5 (5)	4 (6)	3 (10)
vroeggeboorte < 37 weken	1 (2)	3 (4)	5 (5)	4 (6)	2 (7)
totaal	42 (100)	83 (100)	106 (100)	66 (100)	29 (100)
Hematocriet (%) <u>n = 158*</u>	<34	35-38	39-42	>43	
n (%)	n (%)	n (%)	n (%)	n (%)	
geboortegewicht < 2500 gram	0	2 (3)	8 (11)	1 (4)	
vroeggeboorte < 37 weken	0	3 (5)	6 (8)	1 (4)	
totaal	9	59 (100)	72 (100)	22 (100)	

\* De hematocriet was bij de eerste controle slechts bij ongeveer de helft van de vrouwen aangevraagd.

Tabel 6. Relatie tussen Hb respectievelijk Ht bij de eerste zwangerschapscontrole en de verandering van deze parameters tot de dertigste week amenorroe.

		Hemoglobine (mmol/l)					
		<6.9	7.0-7.4	7.5-7.9	8.0-8.4	8.5-8.9	>9.0
gemiddelde							
Hb-verandering	+0.4	-0.3	-0.5	-0.9	-1.2	-1.8	
aantal zwangeren	2	40	83	106	66	29	
		Hematocriet (%)					
		<33%	34-35%	36-37%	38-39%	40-41%	42-43%
gemiddelde							
Ht-verandering	+0.2	-2.5	-4.3	-6.0	-6.4	-7.2	-9.2
aantal zwangeren	5	16	34	36	35	22	9
							6

De mate en de richting van de verandering van Hb en Ht gedurende de zwangerschap hingen sterk samen met de aanvangswaarden (tabel 6). Dit patroon bestond zowel bij de vrouwen die staaltabletten kregen als bij degenen die deze niet hadden gekregen, hoewel in de eerste groepen de stijging in de groep met zeer lage uitgangswaarden meer uitgesproken was.

#### Beschouwing.

In ons onderzoek werd een samenhang gevonden tussen enerzijds een hoog hemoglobinegehalte (8.0 mmol/l en hoger) en een hoge hematocriet (39% en hoger) gemeten bij dertig weken amenorroe en anderzijds een laag geboortegewicht en een korte zwangerschapsduur. Dit verband blijft aanwezig, als gecorrigeerd wordt voor het (aanzienlijke) effect van roken, de (geringe) invloed van pariteit en bloeddruk, en het gebruik van staaltabletten. Het verband is aanwezig in zowel de populatie zwangeren die primair door verloskundigen worden begeleid als onder degenen die vanaf het begin door de gynaecoloog worden gecontroleerd. Een verband tussen het Hb en de Ht bij de eerste controle en de zwangerschapsuitkomst was niet duidelijk. Wel bleek dat de mate van verandering van Hb en Ht sterk samenhangt met de eerst gemeten waarde. Het lijkt er daarbij op dat het lichaam streeft naar een hemoglobinegehalte van rond 7.2 à 7.3 mmol/l en een hematocriet tussen rond de 33 à 34%, later in de zwangerschap. Een dergelijke streefwaarde sluit goed aan bij de bevindingen van de hier gerapporteerde studie en ander onderzoek (1, 2). Immers, vanaf circa 8.0 mmol/l respectievelijk 39% is er meer kans op prematuritas en een laag geboortegewicht en op een secundaire medische indicatie. Bij Hb-waarden onder circa 7.0 mmol/l is er een (bescheiden) samenhang met enkele subjectieve klachten zoals moeheid - erger dan vóór de zwangerschap - en hartkloppingen (hoofdstuk 8).

Over het algemeen stemmen onze bevindingen overeen met de conclusies van anderen: een niet optimale hemodilutie is een risicofactor voor een minder gunstig zwangerschapsbeloop (1 - 10). In onze studie hebben wij kunnen aantonen dat de invloed van deze factor, gemeten via de hematocriet, ook na controle voor een serie relevante, deels met de hematologische parameters en/of zwangerschapsuitkomst geassocieerde, covariabelen aantoonbaar blijft. Het is dan ook niet waarschijnlijk,



dat het pathogenetisch mechanisme door de samenhang met de bloeddruk wordt verklaard, zoals Murphy et al. suggereren (2). Wat roken betreft: ook anderen constateerden, dat het verband tussen hematologische parameters en zwangerschapsuitkomst bij rooksters en niet-rooksters op dezelfde wijze was terug te vinden (1, 2, 4). Het is daarbij wel zo, dat een zwangere die rookt én een hoog Hb heeft een dubbel extra risico loopt.

In tegenstelling tot Murphy et al. (2) vinden wij het verband dat bij dertig weken wordt waargenomen, niet duidelijk terug voor de hematologische bepalingen bij de eerste zwangerschapscontrole. Onze data pleiten voor de hypothese dat niet zozeer de beginwaarden van belang zijn, maar de vraag of men een optimaal traject bereikt. In combinatie met de waarneming van diverse andere groepen dat de verbanden reeds in het tweede trimester duidelijk zijn (4, 5), zou dit optimale traject bij voorkeur al vrij spoedig bereikt moeten zijn.

Het feit dat het verband duidelijker is wat betreft de hematocriet dan wat betreft het hemoglobine, is in overeenstemming met het door Huisman veronderstelde pathofysiologische mechanisme: een hogere viscositeit van het bloed brengt een minder goede placenta-doorbloeding met zich mee (11, 12), en het is dan van belang dat het bloed zich in voldoende mate verdunt. Omdat roken, hoge bloeddruk, alcoholgebruik en pariteit de samenhang tussen hoge Ht- en Hb-waarden en een ongunstige uitkomst niet verklaren, neigen wij ertoe om aan een te geringe hemodilutie een directe causale invloed toe te schrijven. Het ligt voor de hand dat de medicus practicus bij de hematologische controle van de zwangere de Ht altijd (mee)bepaalt. Dit zou, eventueel in samenhang met de bepaling van de bloedviscositeit (11) eenvoudige informatie kunnen geven over de kans op intra-uteriene groeivertraging c.q. aanleiding kunnen zijn om te waken voor (verdere) hemoconcentratie (6). Mede op grond van het optreden van "anemische" klachten zou de optimale range (zie hoofdstuk 8) tussen de 32 en 38% gesitueerd zou kunnen worden. Voor routinematige ijzersuppletie én voor routinematig niet suppleren lijkt geen grond te bestaan.

In een nader onderzoek bij nog eens 400 zwangeren zullen de hier beschreven resultaten worden getoetst. Tevens zal de totale groep dan groot genoeg zijn om meer duidelijkheid te verkrijgen over de invloed van de Hb- en Ht- waarden bij de eerste zwangerschapscontrole.

LITERATUUR.

1. Garn SM, Ridella SA, Petzold AS, Falkner F. Maternal hematologic levels and pregnancy outcomes. *Seminars in perinatology* 1981; 5: 155-162.
2. Murphy JF, O'Riordan J, Newcombe RG, Coles EC, Pearson JF. Relation of haemoglobin levels in first and second trimester to outcome of pregnancy. *The Lancet* 1986; 1: 992-994.
3. Huisjes H. Het hemoglobinegehalte in de zwangerschap. *Nederlands Tijdschrift voor Geneeskunde* 1985; 129: 1323-1324.
4. Huisman A, Aarnoudse JG. Increased 2nd trimester hemoglobin concentration in pregnancies later complicated by hypertension and growth retardation. *Acta Obstetrica et Gynaecologica Scandinavica*. In press.
5. Koller O, Sagen N, Ulstein M, Vaula D. Fetal growth retardation associated with inadequate haemodilution in otherwise uncomplicated pregnancy. *Acta Obstetrica et Gynaecologica Scandinavica* 1979; 58: 9-13.
6. Koller O, Sandvei R, Sagen N. High hemoglobin levels during pregnancy and fetal risk. *International Journal of Gynaecology and Obstetrics* 1980; 18: 53-56.
7. Hytten FE, Paintin DB. Increase in plasma volume during normal pregnancy. *The Journal of Obstetrics and Gynaecology of the British Commonwealth* 1963; 70: 402-407.
8. Gibson HM. Plasma Volume and glomerular filtration rate in pregnancy and their relation to differences in fetal growth. *The Journal of Obstetrics and Gynaecology of the British Commonwealth* 1973; 80: 1067-1074.
9. Pirani BBK, Campbell DM, MacGillivray I. Plasma volume in normal first pregnancy. *The Journal of Obstetrics and Gynaecology of the British Commonwealth* 1973; 80: 884-887.
10. Gallery EDM, Hunyor SN, Györy. Plasma volume concentration: a significant factor in both pregnancy-associated hypertension (pre-eclampsia) and chronic hypertension in pregnancy. *Quarterly Journal of Medicine* 1979; XLVIII; 192: 593-602.
11. Huisman A. De veranderingen in de viscositeit van bloed bij zwangere met intra-uteriene groeivertraging, Groningen 1983.
12. Huisman A. Haemorheological changes during pregnancy. Thesis, R.U. Groningen 1986.

## Hoofdstuk 23

### HEMOGLOBINE EN HEMATOCRIET ALS RISICOFACTOREN VOOR HET KRIJGEN VAN EEN MYOCARDINFARCT.

- een studie in een arbeidsgeneeskundige setting.

J.A. Knottnerus 1) - G.M.H. Swaen 2) - J.J.M. Slangen 2) - A. Volovics 3) - J. Durinck 4).

#### Samenvatting.

De vraag, in hoeverre verhoogde hemoglobine- (Hb-) en hematocriet (Ht-) waarden risicofactoren zijn voor het krijgen van een hartinfarct, is nog niet beantwoord. Met behulp van de gegevens van het Algemeen Periodiek Onderzoek (APO) van Philips' Gloeilampenfabriek werd een patient-controle onderzoek verricht. Patienten waren mannelijke werknemers van 50 tot 60 jaar oud, die (gemiddeld 16 maanden) vóór het krijgen van een eerste infarct een APO hadden ondergaan (n=104). Per patient werden twee gezonde mannelijke controlepersonen getrokken uit de lijst van APO-deelnemers, gematcht op leeftijd. Van patienten en controlepersonen werden de APO-gegevens genoteerd betreffende Hb, Ht, rookgedrag, bloeddruk, cholesterol, polsfrequentie, gewicht, lengte, longfunctie (FEV<sub>5</sub>) en het gebruik van antihypertensieve geneesmiddelen. Na het dichotomiseren van het Hb en de Ht op hun gemiddelde waarden (respectievelijk 9.7 mmol/l en 46%) werden ruwe odds ratio's gevonden van 2.1 (95% betrouwbaarheidsinterval 1.2 - 3.6) en 2.7 (1.6 - 4.6) voor respectievelijk hoge Hb- en Ht- waarden. De verbanden tussen Hb en Ht en het optreden van een myocardinfarct waren ook duidelijk na correctie voor mogelijke versturende factoren door middel van multiële logistische regressie. Onze resultaten zijn in overeenstemming met de hypothese dat matig hoge Ht- en, in mindere mate, Hb-waarden risicofactoren zijn voor het krijgen van een hartinfarct.

- 1) Capaciteitsgroep Epidemiologie, Rijksuniversiteit Limburg
- 2) Capaciteitsgroep Arbeidsgeneeskunde, Rijksuniversiteit Limburg
- 3) Capaciteitsgroep Medische Informatica en Statistiek, Rijksuniversiteit Limburg
- 4) Philips Medische Dienst, Eindhoven.

## Inleiding.

In de medische praktijk worden dagelijks hemoglobine (Hb) bepalingen verricht. Men is daarbij meestal geïnteresseerd in zeer lage waarden, en soms in zeer hoge waarden, als indicatoren van pathologische veranderingen. Veel minder aandacht is er voor minder extreme uitslagen. In dit artikel gaat het om de vraag of er een verband is tussen matig hoge Hb- en hematocriet (Ht)- waarden en de kans op het krijgen van een eerste hartinfarct.

De literatuur geeft over deze vraag geen uitsluitend.

In een aantal prospectieve cohortstudies met lange follow-up perioden werd geen relevant verband gevonden tussen Hb (1-4), Ht (3,5-8) en coronaire hartziekten. In andere studies kwam een dergelijk verband echter wel naar voren, vooral met betrekking tot cardiovasculaire sterfte onder vrouwen (9-13).

Enkele transversale onderzoeken lieten een hogere hematocriet zien bij patiënten met een hartinfarct (14-16) of andere coronaire hartziekten (16), dan bij controlepatiënten. Hoewel deze resultaten niet werden gereproduceerd in twee andere studies (17,18) werden, meer recent, wederom hogere Ht- waarden gevonden bij patiënten met angina pectoris (19) en vernauwingen van de coronaire arteriën (20).

Een nadeel van de transversale methode is dat geen informatie wordt verkregen over de vraag of een hoge Ht voorafgaat aan het infarct of daarvan het gevolg is. Dit laatste zou een uiting kunnen zijn van een compensatiemechanisme, teneinde het ischemische hart van extra zuurstof te kunnen voorzien (3). Voorts is de mogelijkheid geopperd, dat de resultaten van patient-controle onderzoeken in ziekenhuizen vertekend zijn, omdat een verschillende duur van immobilisatie van patiënten en controlepersonen op zichzelf verschillende waarden van de hematologische parameters tot gevolg kan hebben (18).

In een patient-controle onderzoek bestudeerden wij het verband tussen routinematig bepaalde Hb- en Ht- waarden van mannelijke werknemers zonder voorgeschiedenis van coronaire hartziekte en het later optreden van een eerste hartinfarct, rekening houdend met andere bekende risicofactoren.

## Methoden.

### Gegevens bestand.

Alle werknemers van Philips' Gloeilampenfabriek in Eindhoven worden, vanaf de leeftijd van 50 jaar, elke twee jaar uitgenodigd voor een Algemeen Periodiek Onderzoek (APO). Van 1977 tot 1984 hebben ongeveer 10.000 mannelijke werknemers deelgenomen, gemiddeld 3 keer per persoon. Het jaarlijkse responspercentage lag rond de 90%. Alle APO-resultaten zijn opgeslagen in een gecomputeriseerd gegevensbestand.

Het APO bestaat uit:

- een standaardvragenlijst over de gezondheidstoestand, de medische consumptie en het gezondheidsgedrag;
- meting van gewicht, lengte en bloeddruk;
- bloedonderzoek;
- (rust)electrocardiografie en longfunctie-onderzoek.

### Observaties.

De vragenlijsten zijn routinematig afgenomen door speciaal getrainde assistenten. Gevraagd is naar roken ( op het moment en in het verleden), de aanwezigheid van cardiovasculaire klachten of afwijkingen, en het gebruik van antihypertensieve geneesmiddelen.

De bloeddruk is gemeten met een kwik sphygmomanometer met een standaardmanchet om de linkerarm van de onderzochte, in rustig zittende positie. Voor de diastolische bloeddruk is het Korotkov-IV criterium gebruikt.

Lengte en gewicht zijn gemeten nadat schoenen en jasje waren uitgetrokken, en tijdens het tellen van de polsfrequentie lag de onderzochte in rugligging.

Het bloed is afgenomen via venapunctie. In een deel ervan zijn, na behandeling met EDTA, onmiddellijk de hematologische parameters (Hb, Ht, erythrocytengetal (ery's)) bepaald met een TOA-counter CC108. Het cholesterol is bepaald met een Cobas Bio analyzer (Roche), met behulp van cholesteroxidase en -esterase.

Van iedere werknemer is een drie-kanaals rust-electrocardiogram ver

vaardigd, en dit is blind beoordeeld door een ervaren cardioloog op grond van algemeen aanvaarde criteria voor het vaststellen van een hartinfarct (typische infarct-Q-top en ST- en T-afwijkingen).

De longfunctie is bepaald met een droge spirometer (Vica-test), bij geforceerde expiratie. Voor de analyse is de  $FEV_5$  (5 seconden waarde) gebruikt, als de hoogste waarde van drie opeenvolgende metingen.

Selectie van patienten en controlepersonen.

Uit het gegevensbestand werden die mannelijke werknemers geselecteerd die in het kader van het APO hadden gemeld dat zij voor het eerst een hartinfarct hadden gehad. Selectiecriteria waren voorts:

- dat minstens éénmaal eerder een APO was bezocht;
- dat bij vorig(e) bezoek(en) noch anamnestiche, noch electrocardiografische aanwijzingen voor coronaire hartziekte waren gevonden;
- dat het infarct bevestigd werd door vergelijking van de electrocardiogrammen vóór en na het melden van het infarct en/of door informatie van het ziekenhuis waar de patient behandeld was. Deze laatste informatie werd verkregen via de aan het bedrijf verbonden verzekeringsgeneeskundige.

Voor iedere patient werden als controlepersonen at random twee mannelijke werknemers getrokken uit de APO-deelnemerslijst van dezelfde dag dat de patient het infarct meldde, gematcht op geboortjaar. Als een controlepersoon klinisch relevante veranderingen op het ECG had, of geen vorig APO bezocht had, werd hij vervangen door een andere op dezelfde wijze geselecteerde controlepersoon.

Na selectie van patienten en controlepersonen werd op een standaardformulier informatie genoteerd over de hematologische parameters (Hb, Ht, ery's), andere mogelijke determinanten van hartziekte (roken, cholesterol, bloeddruk, polsfrequentie, gewicht, lengte,  $FEV_5$ ) en eventueel gebruik van antihypertensieve geneesmiddelen. Deze informatie werd voor de hartinfarctpatienten ontleend aan het laatste APO vóór het infarct, en voor de controlepersonen werd het daarmee corresponderende APO genomen. In de analyse werd betrokken het rookgedrag (al dan niet roken) ten tijde van dit voorgaande APO.

Statistische analyse.

Ruwe verbanden, niet gecorrigeerd voor de invloed van covariabelen, werden berekend als odds-ratio's met behulp van Schlesselman's methode voor gematchte tripletten (21). Odds-ratio's voor continue variabelen, werden berekend na dichotomisering in "lage" (lager dan of gelijk aan het gemiddelde) en "hoge" (hoger dan het gemiddelde) waarden. Vijfennegentig procents betrouwbaarheidsintervallen werden berekend, ook volgens Schlesselman (21).

Om te controleren voor de andere risico-factoren werd een conditionele logistische regressie-analyse voor gematchte gegevens uitgevoerd volgens Breslow and Day, met het programma PASCAL van het IARC te Lyon (22). De gematchte factoren waren, zoals beschreven, het APO en de leeftijd. De matching ratio was 1:2. In deze analyse werden de continue variabelen met hun gemeten waarde opgenomen, gecentreerd door subtractie van de gemiddelden.

Bij het fitten van het logistisch regressiemodel werd de volgende strategie gevolgd:

- 1) Alle hoofd-effecten werden in het model opgenomen.
- 2) Alle mogelijke producttermen (interacties) van de eerste orde werden gescreend door middel van een voorwaartse selectieprocedure waarbij voorrang werd gegeven aan interacties met de primair bestudeerde risicofactoren (Hb, Ht, en ery's). Hierbij werd een 5% significantiedrempel gehanteerd.
- 3) Niet-significante hoofd-effecten, die niet betrokken waren in interacties werden geëlimineerd onder behoud van de validiteit, dat wil zeggen dat de coëfficiënten van de termen die in het model bleven niet substantieel veranderden in waarde en/of teken.

Er was sprake van multicollineariteit tussen Hb, Ht en ery's. Daarom werd voor elk van deze factoren een apart model opgesteld volgens de beschreven strategie.

Tabel 1: Hematologische parameters en covariabelen bij patiënten (16 + 11 maanden vóór het optreden van het infarct) en controlepersonen, en gemiddelden voor patiënten en controles tezamen.

	Patiënten (n=104)	Controles (n=208)	Allen	
	% gemiddeld	SD	% gemiddeld	SD
Hemoglobine (mmol/l)	9,9	0,7	9,6	0,6
Hematocriet (l/l)	0,47	0,03	0,46	0,03
Erythrocytental (T/l)	5,0	0,4	5,0	0,4
Systolische bloeddruk (mmHg)	135	19	133	17
Diastolische bloeddruk (mmHg)	85	14	84	12
Cholesterol (mmol/l)	6,8	1,1	6,2	1,1
Gewicht (kg)	75,5	8,4	76,9	9,2
Lengte (cm)	173,3	6,3	175,5	6,6
Polsfrequentie (/60 sec)	68	11	69	11
FEV <sub>5</sub> (l)	4,2	0,7	4,4	0,7
Gebruik van antihypertensiva	22	6		11
Roken bij voorgaande APO	85	57		66



## Resultaten.

Honderdvier hartinfarctpatienten en 208 controlepersonen voldeden aan de criteria. de gemiddelde leeftijd van beide groepen was 53 jaar, met een range van 50 tot 60 jaar. De gemiddelde periode tussen het APO waarbij de patienten en controles geselecteerd werden en het voorgaande APO was 31 maanden (SD:13) voor de patienten en 29 maanden (SD:10) voor de controles. De patienten kregen hun infarct gemiddeld 16 maanden na het voorgaande APO met een range van 1 tot 48 maanden.

Tabel 1 geeft informatie over de waarden van hematologische parameters en covariabelen ten tijde van het voorgaande APO. Hartinfarctpatienten lijken wat hogere Hb- en Ht- waarden te hebben gehad dan de controles, wat nog duidelijker gedemonstreerd wordt in tabel 2, waarin de verdelingen van het Hb en de Ht in beide groepen worden gegeven.

Tabel 2: Distributie van hemoglobine en hematocriet bij patienten (vóór het optreden van het infarct) en controles.

HEMOGLOBINE						HEMATOCRIET					
<u>Patienten</u>						<u>Controles</u>					
Hb	n	%	n	%		Ht	n	%	n	%	
(mmol/l)						(%)					
8,0 -	1	1,0	5	2,4		36 -	0	0,0	1	0,5	
8,5 -	5	4,8	20	9,6		38 -	0	0,0	1	0,5	
9,0 -	26	25,0	58	27,9		40 -	2	1,9	9	4,3	
9,5 -	28	26,9	68	32,7		42 -	8	7,7	34	16,3	
10,0 -	23	22,1	39	18,8		44 -	19	18,3	63	30,3	
10,5 -	16	15,4	16	7,7		46 -	32	30,8	46	22,1	
11,0 -	3	2,9	2	1,0		48 -	19	18,3	38	18,3	
11,5 +	2	1,9	0	0,0		50 -	17	16,3	13	6,3	
						52 -	4	3,8	3	1,4	
						54 -	2	1,9	0	0,0	
						56 +	1	1,0	0	0,0	
	104 (100)		208 (100)				104 (100)		208 (100)		

De hartinfarctpatienten hebben meer gerookt en hadden tevoren ook een hoger cholesterolgehalte dan de controles (tabel 1). Bijna een kwart van de patienten en 6% van de controles meldden op het voorafgaande APO antihypertensiva te gebruiken. Waarschijnlijk waren deze middelen voorgeschreven door hun huisarts, en dit kan ertoe hebben bijgedragen

dat tussen de groepen geen verschillen werden gevonden in systolische en diastolische bloeddruk.

In tabel 3 worden de ongecorrigeerde odds ratio's gegevens van hematocriet en hemoglobine, en van de covariabelen met een odds ratio significant groter dan 1.0, alsmede van de bloeddruk, na dichotomisering op het gemiddelde weergegeven. Voor het Hb (afkappunt 9.7 mmol/l) en de Ht (afkappunt 0.46) werden odds ratio's gevonden van respectievelijk 2.1 en 2.7, significant groter dan 1.0. Zoals verwacht vertonen ook roken en cholesterolgehalte duidelijk positieve relaties met het optreden van een infarct. De odds ratio van de bloeddruk verschilt niet significant van 1.0. Zoals gezegd, kan dit het gevolg zijn van medicatie. Aangezien antihypertensiva met name worden voorgeschreven bij vastgestelde hoge bloeddruk, kan voor hun effect in de analyse niet meer gecorrigeerd worden. Het wekt in dit licht overigens geen verwondering, dat het gebruik van antihypertensiva geassocieerd is met een aanzienlijk verhoogd risico op hartinfarct.

Tabel 3: Ongecorrigeerde Odds Ratio's voor het risico op myocardiinfarct (niet gecorrigeerd voor covariabelen). Continue variabelen gedichotomiseerd op het gemiddelde in "hoog" (hoger dan het gemiddelde) versus "laag" (gelijk aan of lager dan het gemiddelde).

	OR	95% CI
Hemoglobine	2,10	1,22 - 3,61
Hematocriet	2,72	1,59 - 4,64
Roken (actueel)	4,87	2,56 - 9,22
Cholesterol	2,87	1,64 - 4,69
Antihypertensiva gebruik	5,25	2,23 - 11,13
Bloeddruk *	1,12	0,61 - 2,06

\* Bloeddruk = diastolische druk + 1/3.(systolische - diastolische druk).

Als resultaat van de multiële logistische regressie worden de volgende modellen gepresenteerd (tabel 4): alle variabelen zonder interactietermen (1), en de modellen voor de hematocriet (2), het hemoglobine (3) en het erythrocytengetal (4), telkens met andere risicofactoren die een significant effect vertonen en een gevonden interactie tussen de hematologische parameters en de lengte. De hematocriet blijkt een sterk effect te hebben na correctie voor de andere factoren. Het effect van het Hb is geringer en niet significant in het volledige

model. Worden Hb en Ht in aparte modellen ondergebracht, dan zien we dat het effect van het Hb significant wordt. Een invloed van het erythrocytengetal kan niet worden aangetoond.

Na de correctie voor andere factoren wordt het risico dat samenhangt het gebruik van medicatie nog groter, vergeleken met de ruwe odds ratio ( $e^{2.19} = 8.9$ , versus 5.3), en blijven de effecten van roken en cholesterol duidelijk aantoonbaar. Verschillen in bloeddruk en FEV<sub>5</sub> hebben geen aantoonbare samenhang met het risico. De Pearson correlatiecoëfficiënt tussen lichaamsgewicht en lengte bedroeg 0.58, maar er was geen significant effect van het gewicht aantoonbaar noch was het betrokken in een interactie, ook als lichaamslengte buiten de vergelijking werd gehouden. Interacties werden gevonden tussen lengte en de hematologische parameters. Hoewel niet significant volgens het 5% criterium is de interactie tussen lengte en erythrocyten in model (4) opgenomen ter vergelijking met de modellen (2) en (3).

Om een indruk te geven van de effecten van Ht en Hb: de odds ratio voor de vergelijking van een man met een gemiddelde lengte en een Ht van 0.49 (= gemiddelde + standaarddeviatie) met een overigens vergelijkbare man met een Ht van 0.43 (= gemiddelde - standaarddeviatie) is gelijk aan:  $\exp(31.19(0.49-0.43)) = 6.5$ . De odds ratio van een man met een gemiddelde lengte en een Hb van 10.3 (= gemiddelde + standaarddeviatie) ten opzichte van een man met een Hb van 9.1 (= gemiddelde - standaarddeviatie) is gelijk aan:  $\exp(0.91(10.3-9.1)) = 3.0$ .

#### Discussie.

Onze resultaten wijzen op een positief verband tussen matig hoge Ht-en, in mindere mate, Hb-waarden enerzijds en het risico op een hartinfarct anderzijds. Hierbij werd uitgegaan van Ht- en Hb-bepalingen die reeds vóór het infarct verricht waren. De verbanden waren ook duidelijk aantoonbaar na correctie voor roken, cholesterol, bloeddruk, lichaamsgewicht, lichaamslengte en longfunctie. Deze bevindingen zijn in overeenstemming met de resultaten van Campbell et al, hoewel de laatste nog geen gegevens over rookgedrag in de analyse kon betrekken (11). Omdat de APO-gegevens verzameld en geregistreerd waren vóórdat de studie werd verricht, kunnen noch de beantwoording van de vragenlijst, noch het aflezen en coderen van de meetgegevens beïnvloed zijn

Tabel 4: Multipele logistische regressie modellen voor het risico op hartinfarct, met de continue variabelen gecentreerd op hun gemiddelde \*

	(1)		(2)		(3)		(4)				
	Beta	SE	Beta	SE	Beta	SE	Beta	SE	Gemidd.	SD	Eenheid
Hematocriet (Ht)	38,44	11,31	31,19	6,88					0,46	0,03	ml/ml
Hemoglobine (Hb)	0,11	0,51			0,91	0,28			9,7	0,6	mmol/l
Erythrocytengetal (Er)	-0,84	0,70					0,78	0,43	5,0	0,4	T/l
Roken (actueel)	1,54	0,44	1,88	0,46	1,87	0,44	1,84	0,42			
Cholesterol	0,75	0,18	0,76	0,18	0,74	0,17	0,65	0,16	6,4	1,1	mmol/l
Antihypertensiva gebruik	2,19	0,60	2,20	0,58	2,04	0,54	1,97	0,51			
Gewicht	0,03	0,03							76,4	9,0	kg
Lengte (Le)	-0,07	0,04	-0,06	0,03	-0,06	0,03	-0,05	0,02	174,8	6,6	cm
Bloeddruk **	0,01	0,01							101	13	mmHg
FEV <sub>5</sub>	0,04	0,26							4,35	0,74	l
Polsfrequentie	-0,01	0,02							69	11	min <sup>-1</sup>
(Ht-Ht) (Le-Le)			2,51	0,899							
(Hb-Hb) (Le-Le)					0,14	0,066					
(Er-Er) (Le-Le)							0,12	0,064			

\* in de regressievergelijking wordt dus van een continue variabele altijd eerst het gemiddelde afgetrokken.

\*\* Bloeddruk = diastolische druk + 1/3.(systolische - diastolische druk)

door voorkennis over (het doel van) het onderzoek. Bovendien kan de aanwezigheid van het infarct de gebruikte APO-gegevens over de risicofactoren niet beïnvloed hebben omdat deze al vóór het optreden van het infarct vastgelegd waren. In ons onderzoek was ook geen vertekening mogelijk door verschillen in hospitalisatie tussen patienten en controlepersonen (18) omdat allen ten tijde van de gebruikte APO-metingen goed gezond en ambulant waren.

Vertekening kan wel veroorzaakt zijn doordat personen met een fataal infarct, of degenen die daarna het APO niet meer bezochten (bijvoorbeeld vanwege invaliditeit), niet in het onderzoek werden betrokken. Dit heeft mogelijk geleid tot een "bias to the null", omdat in cohort-studies juist effecten van Hb en Ht ten aanzien van cardiovasculaire mortaliteit zijn gevonden (9-13).

Voor de gevonden verbanden tussen Ht, Hb en hartinfarct komen verschillende verklaringen in aanmerking. Waarschijnlijk is, dat de viscositeit van het bloed een rol speelt. De positieve invloed van de hematocriet op de viscositeit is bekend (19,23), evenals het negatieve verband tussen de bloedviscositeit enerzijds en de coronaire doorbloeding (23) respectievelijk cardiale efficiency anderzijds (24,25). Analoge verbanden zijn gevonden tussen de cerebrale doorbloeding en een hoge Ht ( $> 0.46$ ) (26,27). Een hogere Ht is ook geassocieerd met ernstiger graden van coronaire stenose (20) en van cerebrale infarctering (28). Het feit dat wij een duidelijke invloed van de Ht vinden, en niet van het erythrocytengetal, houdt in dat het volume van de rode cel het risico positief beïnvloedt.

In postmortale studies is een toegenomen aantal intercoronaire anastomosen gevonden bij anemische patienten, wellicht als compensatie van de minder goede zuurstofbindende capaciteit van het bloed (29). Dit zou anemische patienten kunnen beschermen tegen de gevolgen van coronaire trombose (30). In diverse studies is een negatief verband tussen Ht en bloedingstijd waargenomen. Deze bevinding wordt mogelijk verklaard doordat een hoge Ht de snelheid van de bloedstroom vertraagt en aldus plaatjesaggregatie en tromboseneiging bevordert (31,32). Een hoog gemiddeld celvolume kan dit proces versterken.

Positieve verbanden zijn gevonden tussen de Ht en het Hb en bekende risicofactoren zoals roken (9,18,33,34), cholesterol- en triglycerideconcentratie (35-37), en bloeddruk (18,38).

Wat betreft roken zou er sprake kunnen zijn van een compensatie mechanisme voor het hogere Hb-CO gehalte. De effecten van Ht en Hb worden in onze studie echter niet door roken verklaard, omdat zij na correctie voor roken blijven bestaan.

Gesuggereerd is dat een toename van de hematocriet, en dus een vermindering van het relatieve plasmaplasma volume, een stijging van het cholesterolgehalte veroorzaakt (35-37). Aldus zou het atherogene effect van cholesterol voor een deel het indirecte gevolg kunnen zijn van een hoge Ht. Wij vonden echter nauwelijks een verband tussen Ht en cholesterolgehalte ( $r=0.05$ ). Bovendien verdwenen de relaties tussen Ht, Hb en hartinfarct niet na controle voor het cholesterolgehalte.

De correlatiecoëfficiënt tussen bloeddruk en Ht was slechts 0.21, en in ons onderzoek was bloeddruk geen significante risicofactor. Het is dus niet waarschijnlijk dat hogere bloeddruk het risico, dat met een hogere Ht geassocieerd is, verklaart. We moeten hierbij voorzichtig zijn omdat ten aanzien van de relatie bloeddruk-hartinfarct "confounding by indication" wat betreft het voorschrijven van antihypertensiva kan zijn opgetreden. Ook het verband tussen Ht en bloeddruk kan door deze middelen negatief zijn beïnvloed. Uitgesloten moet evenwel niet worden geacht, dat de predictieve waarde van hoge bloeddruk als risicofactor voor ischemische hartziekten boven de leeftijd van 50 jaar sterk afneemt (39).

Wij vonden wat betreft het risico op een hartinfarct een interactie tussen Ht en Hb enerzijds en lichaamslengte anderzijds. Of dit een biologische betekenis heeft is ons niet duidelijk. Voor de beantwoording van deze vraag zou nader onderzoek nodig zijn.

Een helder en aanvaardbaar concept van het mechanisme, verantwoordelijk voor het positieve verband tussen Ht, Hb en hartinfarct, is nog niet uitgekristalliseerd. Een protopathisch effect (40) - een toename van de Hb synthese als gevolg van langdurende ischemie - kan niet worden uitgesloten. Maar, gezien de negatieve cardiale anamnese én de niet voor ischemie verdachte electrocardiogrammen van onze patienten en controles ten tijde van de bloedbepalingen, is dit geen voor de

hand liggende verklaring. Mede gegeven de resultaten van de cohortstudie van Campbell et al, wijzend op een constant risico gedurende de follow-up periode (13), lijkt een causaal verband tussen hoge Ht- en Hb- waarden en het optreden van een hartinfarct aannemelijk, al dan niet gesuperponeerd op of gepotenteerd door subklinische ischemie. Beide mechanismen (een protopathisch en een causaal effect) kunnen ook tegelijk bestaan en elkaar versterken.

Gegeven de bevindingen van het biomedisch onderzoek en de nog uiteenlopende resultaten van het epidemiologisch onderzoek, is nader onderzoek naar de precieze aard van het verband tussen hematologische parameters en coronaire hartziekte gewenst, mede met het oog op mogelijke klinische consequenties.

#### LITERATUUR.

1. Paul O, Lepper MH, Phelan WH et al. A longitudinal study of coronary heart disease. *Circulation* 1963; 28: 20-31.
2. Truett J, Cornfield J, Kannel W. A multivariate analysis of coronary heart disease in Framingham. *J Chron Dis* 1967; 20: 511-24.
3. Abu-Zeid HAH. The relation between haemoglobin level and the risk for ischemic heart disease: a prospective study. *J Chron Dis* 1976; 29: 395-403.
4. Yano K, Reed DM, McGee DL. Ten-year incidence of coronary heart disease in the Honolulu heart program. Relationship to biologic and lifestyle characteristics. *Am J Epidemiol* 1984; 119: 653-66.
5. Wilhelmsen L, Wedel H, Tibblin G. Multivariate analysis of risk factors for coronary heart disease. *Circulation* 1973; 48: 950-58.
6. Takkunen H, Aromaa A. Mortality and anaemia (Letter to the editor). *The Lancet* 1974; 2: 523-24.
7. Brand RJ, Rosenman RH, Sholtz RI, Friedman M. Multivariate prediction of coronary heart disease in the Western Collaborative Group Study compared to the findings of the Framingham study. *Circulation* 1976; 53: 348-55.
8. Szatrowski TP, Peterson AV, Shimizu Y, Prentice RL, Mason MW, Fukunaga Y, Kato Hiroo. Serum Cholesterol, other risk factors and cardiovascular disease in a Japanese cohort. *J Chron Dis* 1984; 37: 569-584.

9. Elwood PC, Waters WE, Benjamin IT, Sweetman PM. Mortality and anaemia in women. *The Lancet* 1974; 2: 891-94.
10. Cullen KJ, Stenhouse NS, Wearne KL. Raised haemoglobin and risk of cardiovascular disease. *The Lancet* 1981; 4: 1288-1289.
11. Campbell MJ. The analysis using GENSTAT of anaemia, sugar intake and Quetelet's Index as prognostic indicators in women. *Statistics in Medicine* 1983; 2: 223-28.
12. Schatzkin A, Cupples A, Heeren T et al. Sudden death in the Framingham Study. Differences in incidence and risk factors by sex and coronary disease status. *Am J Epidemiol* 1984; 120: 888-99.
13. Campbell MJ, Elwood PC, Mackean J, Waters WE. Mortality, haemoglobin level and haematocrit in women. *J Chron Dis* 1985; 38: 881-889.
14. Burch GE, De Pasquale NP. The haematocrit in patients with myocardial infarction. *JAMA* 1962; 180: 63-65.
15. De Pasquale NP, Burch GE. Haematocrit in women with myocardial infarction. *JAMA* 1963; 183: 142-43.
16. Mayer GA. Haematocrit and coronary heart disease. *Canad Med Ass J* 1965; 93: 1151-53.
17. Conley CL, Russell RP, Thomas CB, Tumulty PA. Haematocrit values in coronary artery disease. *Arch Intern med* 1964; 113: 170-76.
18. McDonough JR, Hames CG, Garrison GE et al. The relationship of haematocrit to cardiovascular states of health in the negro and white population of Evans County, georgia. *J Chron Dis* 1965; 18: 243-57.
19. Nicolaides AN, Bowers R, Horbourne T et al. Blood viscosity, red-cell flexibility, haematocrit, and plasma-fibrinogen in patients with angina. *The Lancet* 1977; 2: 943-45.
20. Lowe GDO, Drummond MM, Lorimer et al. Relation between extent of coronary artery disease and blood viscosity. *Br Med J* 1980; 280: 673-74.
21. Schlesselman JJ. Case control studies, 9th ed. Oxford University Press 1982.
22. Breslow NE, Day NE. Statistical methods in cancer research. Vol. 1 The analysis of case-control studies. Lyon: IARC Scientific Publications 1980.
23. Burch GE, De Pasquale NP. Haematocrit, viscosity and coronary blood flow. *Diseases of the Chest* 1965; 48: 225-30.



24. Dintenfass L, Lake B. Blood viscosity factors in evaluation of submaximal work output and cardiac activity in men. *Angiology* 1977; 28: 788-98.
25. Dintenfass L, Lake B. Blood viscosity factors in evaluation of submaximal work output and cardiac activity in patients with myocardial infarction and angina, and in normals. 9th Europ. Conf. Microcirculation, Antwerp 1976. *Bibl Anat* no 16: 498-500 (Karger, Basel 1977).
26. Thomas DJ, Du Boulay GH, Marshall J et al. Effect of haematocrit on cerebral blood-flow in man. *The Lancet* 1977; 2: 941-45.
27. Grotta J, Ackerman R, Correia J et al. Whole blood viscosity parameters and cerebral flow. *Stroke* 1982; 13: 296-301.
28. Harrison MJG, Pollock S, Kendall BE, Marshall J. Effect of haematocrit on carotid stenosis and cerebral infarction. *The Lancet* 1981; 2: 114-5.
29. Zoll PA, Wessler S, Schlesinger J. Interarterial coronary anastomoses in human heart, with particular reference to anemia and relative cardiac anoxia. *Circulation* 1951; 4: 797-815.
30. Elwood PC. Evaluation of the clinical importance of anemia. *The American Journal of Clinical Nutrition* 1973; 26: 958-64.
31. Turitto VT, Weiss HJ. Red blood cells: their dual role in thrombus formation. *Science* 1980; 207: 541-3.
32. Editorial. The bleeding-time and the haematocrit. *The Lancet* 1984; 1: 997-8.
33. Abu-Zeid HAH, Chapman JM. The relationship between haemoglobin level and some risk factors in ischemic heart disease. The Los Angeles Heart Study *Circulation* 1973; 48 (suppl 4): 9.
34. Dintenfass L. Elevation of blood viscosity, aggregation of red cells, haematocrit values and fibrinogen levels in cigarette smokers. *Med J Austr* 1975; 1: 617-20.
35. Elwood PC, Mahler R, Sweetman P et al. Association between circulating haemoglobin level, serum-cholesterol and blood pressure. *The Lancet* 1970; 1: 589-90.
36. Böttiger LE, Carlson LA. Relation between cholesterol and triglyceride concentration and haemoglobin values in non-anaemic healthy persons. *Br med J* 1972; 3: 731-3.
37. Böttiger LE, Carlson LA. Haemoglobin, plasma lipids, and coronary

- heart disease. Am Heart J 1973; 86: 842-4.
38. Harlan WR, Hull AL, Schmouder RL et al. Blood pressure and nutrition in adults. The national health and Nutrition Examination Survey. Am J Epidemiol 1984; 120: 17-28.
39. Barrett-Connor E, Suarez L, Khaw K et al. Ischemic heart disease risk factors after age 50. J Chron Dis 1984; 903-8.
40. Feinstein AR. Clinical Epidemiology. The architecture of clinical research. WB Saunders Co., Philadelphia 1985.



DEEL VI

BESCHOUWING EN SAMENVATTING.

## Hoofdstuk 25

---

### BESCHOUWING.

#### 1. Het werkveld van de klinische epidemiologie.

Als iemand om een bepaalde reden contact heeft gelegd met dokter, dan zijn klachten, symptomen, de resultaten van eventueel verricht diagnostisch onderzoek en de interpretatie van al deze gegevens in hoge mate bepalend voor de verdere 'patiënten carrière'. Daarnaast hebben ook persoonskenmerken, zoals de leeftijd van de hulpvragende en de voorgeschiedenis, invloed wat betreft zowel de waarschijnlijkheid van bepaalde diagnoses als de verder te volgen strategie. De laatste bijvoorbeeld vanwege de mate van belastbaarheid van de patiënt ten aanzien van verdere diagnostiek en therapie. Last but not least zijn van betekenis de wensen en de voorkeuren van de patiënt, en ook de opvattingen, voorkeuren en de werksituatie van de arts (1-5).

Een scala van factoren markeert dus de besluitvorming die leidt tot een al dan niet optimaal gedoseerde medische aandacht voor onze patiënt. "Optimaal gedoseerd" betekent dan zowel een juiste hoeveelheid als een juiste inhoud van die aandacht, en niet-optimaal zowel te veel als te weinig alsmede niet-adequate medische aandacht.

Wat betreft de interactie arts-patiënt dient gestreefd te worden naar een synthese van het 'deskundigheidsmodel' (de arts heeft kennis en vaardigheden tot zijn beschikking) en het 'mondigheidsmodel' (de patiënt heeft het eerste en het laatste woord, en heeft bovendien vragen, ervaringen en ideeën die niet gemist kunnen worden). De voorkeur heeft een benadering waarbij de arts naar aanleiding van de vraag van de patiënt een deskundige beoordeling verricht, en vervolgens zijn diagnostische overwegingen en conclusies en eventuele beleidsalternatieven aan de patiënt voorlegt, gewogen naar waarschijnlijkheid, ernst, mate van belasting, therapeutische consequenties en prognose. De ingangsvraag dient daarbij dan wel voldoende verhelderd te zijn (6), en de patiënt moet over de diverse alternatieven ook kunnen oordelen en beslissen (7).

Wie deze zinnen leest, realiseert zich hoe complex het gebeuren is waarover we spreken. In de eerste plaats leidt dit tot een zekere wetenschappelijke bescheidenheid: de dagtaak van de medicus practicus

omvat een aaneenschakeling van zulke gebeurtenissen, en deze móet er mee uit de voeten, hoe onbevredigend ook de stand van de wetenschap is. Tevens is duidelijk, en dit kan ook als geruststellend worden ervaren, dat de individuele menselijke factor altijd een centraal element zal blijven. Tegelijkertijd liggen er voor uiteenlopende wetenschappelijke disciplines vele aangrijpingspunten, om een bijdrage te leveren aan de verbetering van de voorwaarden op grond waarvan de medische hulpverlening plaatsvindt.

Tot de basisvoorwaarden voor het op zinvolle wijze kunnen bereiken van de genoemde synthese behoren niet alleen het goed omschreven zijn van de deskundigheid van de arts maar ook het valide zijn van het door hem in dat kader gehanteerde pakket aan kennis en vaardigheden. Klinisch-epidemiologische onderzoeken en analyses kunnen met name bijdragen aan dit laatstgenoemde "sine qua non". Het betreft dan zowel het toetsen en schonen van reeds in gebruik zijnde medische kennis en inzichten, als het toetsen en (mede-)ontwikkelen van nieuwe. Deze bijdrage kan liggen op (1) het niveau van het detail, het concrete, het direct-ambachtelijke, maar kan ook (2) conceptueel van aard zijn en van betekenis voor het ontsluiten van nieuwe denkrichtingen. Een andere belangrijke onderscheiding in deze bijdrage betreft (a) het vergaren en toepasbaar maken van kwantitatieve informatie enerzijds, en (b) kwalitatieve analyse anderzijds. De genoemde elementen kan men in onderlinge samenhang beschouwen, en zij leveren dan een schematische indeling op van het werkveld van de klinische epidemiologie (Figuur 1).

In dit proefschrift is, in wisselende mate, ten aanzien van ieder van deze "vier velden van de klinische epidemiologie" materiaal gepresenteerd. Onvermijdelijk was dat bepaalde deelgebieden niet rechtstreeks aan bod kwamen, zoals effectiviteitsonderzoek. Daarop gaan wij elders in (8).

Figuur 1: Het werkveld van de klinische epidemiologie, ondergebracht in een viervelden tabel, met voorbeelden.

	conceptueel	concreet
kwalitatief	<ul style="list-style-type: none"> <li>. selectie ten gevolge van verwijzing</li> <li>. wat is normaal</li> <li>. de beslisboom</li> </ul>	<ul style="list-style-type: none"> <li>. wat is een logisch gefaseerde en hanteerbare strategie bij een klier in de hals</li> </ul>
kwantitatief	<ul style="list-style-type: none"> <li>. kwantificeren van selectiebias</li> <li>. methoden om het optimale afkappunt te bepalen</li> <li>. sensitiviteitsanalyse</li> </ul>	<ul style="list-style-type: none"> <li>. voorspellende waarde van bepaalde klachten</li> <li>. wat is de strategie met de "hoogste verwachte waarde" bij een klier in de hals</li> <li>. effectiviteit van endarterectomie</li> </ul>

## 2. Selectie en selecteren.

Ten aanzien van de validiteit van het door de medicus gehanteerde kennispakket is veel aandacht besteed aan verschillen tussen de patiëntenpopulaties in eerste en tweedelijns geneeskunde, en de rol die verwijzingspatronen hierbij spelen (hoofdstukken 2, 3, 12, 13, 14, 21). Aan bod kwamen verschillen in klachten- en morbiditeitspatronen, de voorspellende waarde van klachten, symptomen en andere diagnostische gegevens, en de mate van ontwikkeling van ziektebeelden. Verschillen zijn er ook te verwachten wat betreft het verloop, de prognose en de gevoeligheid voor therapeutische beïnvloeding (8, 9). Het zich realiseren van deze verschillen is van groot belang. Ten aanzien van de wetenschappelijke bestudering van het 'kennispakket' is er immers nog een forse achterstand van de huisartsgeneeskunde ten opzichte van de specialistische geneeskunde. Het is dan voor de huisarts verleidelijk om bij de specialist 'leentjebuur' te spelen, hetgeen - mede bij gebrek aan alternatieven - bevorderd wordt in het medisch basiscurriculum en bepaalde vormen van nascholing. Dit scholings- en vormingsproces leidt dan weer tot remming van de eigen ontwikkeling en vergroting van de achterstand. Het geschetste cirkelproces heeft tot gevolg, dat huisartsen dikwijls werken met klinische basiskennis, die niet op hun positie en 'populatie' is toegesneden. Zij moeten zelf de nodige aanvullingen trachten te creëren, voornamelijk op grond van de

eigen, weinig gesystematiseerde praktijkervaring. Dit staat in scherp contrast tot de steeds grotere vlucht die de wetenschappelijke onderbouwing van de specialistische geneeskunde neemt over een steeds breder terrein met steeds meer subterreinen. De huisarts is door deze gang van zaken steeds minder in staat om een zelfbewuste gesprekspartner te zijn van de specialist, op grond van een krachtig ontwikkelde eigen discipline.

Een en ander draagt ertoe bij dat het functioneren van de huisarts onder druk staat, wat betreft zowel het eigen algemeen geneeskundig ambacht, als de te vervullen zeef functie naar de specialistische zorg. Een stevige, op dit ambacht gerichte wetenschappelijke basis is immers onontbeerlijk, teneinde te komen tot een optimale selectie van patiënten voor de tweede-lijns gezondheidszorg. Bedacht moet worden dat de zeef functie van de huisarts in belangrijke mate bepalend is voor de kwaliteit en de kosten van de gezondheidszorg.

Ondanks de genoemde druk bestaat de zeef functie van de huisarts nog steeds, en daarmee een belangrijk fundament van de huisartsgeneeskundige identiteit (hoofdstukken 9, 10, 12). Hier is een opmerkelijke paradox te constateren: naarmate de huisarts sterker (voor)selecteert, zijn er grotere verschillen tussen huisartsgeneeskundige en specialistische patiëntenpopulaties te verwachten. De huisartsgeneeskunde kan dit - stank voor dank - op haar brood krijgen in de vorm van een zich sterker van haar identiteit vervreemdend onderwijs én onderzoek vanuit de medische faculteiten met als gevolg juist een ondermijning van haar zeef functie (wellicht is hiermee een stukje geschiedenis van de relatie medische faculteiten - huisartsgeneeskunde kernachtig beschreven). Indien de huisarts niet, of slechts met behulp van een dobbelsteen, zou selecteren, dan zouden de wezenlijke verschillen tussen beide populaties vervagen en zou alle spraakverwarring wegvallen. Hetzelfde zou dan evenwel het lot zijn van de branche, respectievelijk de professionele identiteit van de huisarts. Een ingebouwde beveiliging tegen een dergelijke ontwikkeling is de blijvende behoefte aan een niet over te dragen en vitaal element van de huisartsgeneeskunde: het all-round-karakter, het overzicht. Per definitie kan géén ander specialisme dit overnemen. Wie zal er voorselecteren en richtingaangeven als de wegwijzer wegvalt? Het zou te gemakkelijk zijn te zeggen dat de



mondige patiënt dit zelf wel zou kunnen. Hoe zou, toevalstreffers daargelaten, bijvoorbeeld de 'patiënt carrière' eruit zien van de mondige patiënt met onbegrepen moeheidsklachten die na het lezen van de medische encyclopedie besloten heeft om zich als eerste te wenden tot, laten we zeggen, de endocrinoloog?

Aantrekkelijker is het om aan te sluiten bij de op dit moment nog intacte onderdelen van de zeeffunctie van de huisarts, en deze groten-deels op ervaringskennis berustende functie van een wetenschappelijke en wetenschappelijk erkende basis te voorzien, met onderkenning van de 'kliniek van de huisartsgeneeskunde' (10).

### 3. "Vage klachten".

Het selecterend bezig zijn van de arts wordt sterk op de proef gesteld als de patient zich aandient met zogenaamde vage klachten die niet duidelijk op een bepaald organisch of psychosociaal substraat wijzen. Wat diagnostiek betreft, staat hem ten dienste een veelheid aan middelen en een armoede aan houvast. Welke onderzoeken zijn geïndiceerd naar aanleiding van welke klachten? Om deze vraag te beantwoorden is het allereerst nodig om te constateren dat de vraag helaas niet in één keer te beantwoorden is. Men moet stapsgewijs te werk gaan, en zich in eerste instantie beperken tot één of enkele klachten én een beperkt aantal diagnostica. In de hoofdstukken 4 t/m 8 hebben we dit gedaan. De conclusies lijken dan ook op het eerste gezicht van beperkt belang: hoofdpijnklahten die zodanig ernstig zijn dat men regelmatig hoofdpijnmiddelen moet gebruiken hebben geen relevante betekenis voor het herkennen van hoge bloeddruk, en vormen dus ook geen speciale indicatie voor (extra) bloeddrukmeting, en: moeheidsklachten in de huisartspraktijk, ook degene waarvoor geen andere verklaring voor de hand ligt, vertonen geen verband met hematologische parameters, noch met het serumferritine gehalte. Bij nadere beschouwing kan men echter inzien dat het trekken van de praktische consequenties uit deze observaties een groot aantal routinematige handelingen en hulpvragen op losse schroeven zet, alsmede de ermee gepaard gaande medicalisering en financiële kosten.

Nu zijn er voor het verrichten van handelingen natuurlijk meer redenen dan de overtuiging van de arts dat ze, medisch gezien, zinvol zijn. Vaak is bijvoorbeeld de vraag van de patiënt doorslaggevend, ook al heeft de dokter twijfels, of zelfs een andere opvatting over wat de beste weg is. Los van de vraag wat in zo'n situatie het juiste besluit is, is het van belang in welke mate de arts kan (en durft) terug (te) vallen op vakmatige overtuiging. Dit zal eerder gebeuren, naarmate ervaringskennis meer wordt ondersteund door gepubliceerde onderzoeksgegevens. Ook voor de patiënt is zo'n onderbouwing een aanvaardbaar houvast. Voorts zullen ook anderen, bijvoorbeeld gebruikersgroeperingen, hun voordeel kunnen doen met dergelijk materiaal.

Een gevolg kan ook zijn, dat klachten meer specifiek worden geanalyseerd, op zoek naar de wél en niet informatieve aspecten ervan, en daarmee naar meer gerichte indicatiestellingen voor nadere diagnostiek. Wij hebben getracht wat betreft de klacht moeheid in de huisartspraktijk de 'ruis' te elimineren door alle reeds in het eerste consult voor de hand liggende verklaringen uit te sluiten. De overweging hierbij was, dat wij hierbij optimaal zouden aansluiten bij de mede door ervaring gevormde kennis van de huisartsen. Bij het onderzoek naar het verband tussen moeheid en de hematologie van zwangeren kozen we een andere benadering. Specifieke kenmerken van de moeheidsklacht en andere klachten werden expliciet nagevraagd. Hierbij kwam van een aantal klachten (onder andere moeheid, erger dan vóór de zwangerschap, hartkloppingen, duizeligheid en kuitkrampen) wél enig onderscheidend vermogen naar voren ten aanzien van een laag Hb-gehalte. Een mogelijke verklaring hiervoor is, dat bij het lagere Hb-gehalte van zwangeren wel een verband met klachten bestaat. Omdat de klachten en het Hb-gehalte in beide studies onafhankelijk van elkaar werden vastgelegd achten wij niet waarschijnlijk dat het gebruik van andere methoden de verschillende uitkomsten verklaart. Geconcludeerd kan wel worden dat het aan patienten zelf voorleggen van gedetailleerde vragenlijsten het evalueren van verfijnde nuances van klachten mogelijk maakt.

Een algemene conclusie is, dat de vraag naar de zin van bepaalde diagnostiek naar aanleiding van de - zo frequent voorkomende (11) - vage klachten zich goed leent voor klinisch-epidemiologisch onderzoek.

#### 4. Medische Besluitvorming.

Het handelen van artsen is zinvol voor zover het kan bijdragen tot een verbetering van de prognose van patiënten, en voor zover de voordelen van dat handelen opwegen tegen de nadelen.

Deze eisen gelden al ten aanzien van de eerste fase van de medische betrokkenheid, de diagnostiek. De diagnose is geen doel op zich, maar vindt zijn legitimatie als aangrijpingspunt voor het beleid (12) (hoofdstuk 17). Indien het beleid niét afhangt van het stellen van de bedoelde diagnose, dan kan hierop gericht onderzoek achterwege blijven. Als men met dit principe in het achterhoofd het referentiewaarden-vraagstuk analyseert, dan blijkt dit vraagstuk vloeiend over te gaan in een besliskundige probleemstelling: onder welke omstandigheden is het doen van diagnostiek zinvol, de a priori kansen op ziekte(n) en de eventuele consequenties in termen van beleid en gezondheidswinst in aanmerking genomen? En: hoe moet vervolgens een gegeven uitslag geïnterpreteerd worden? In het kader van deze probleemstelling blijkt onder andere dat het voor tests met een continue schaal van uitslagen beter is om informatie te hebben over de gehele distributie van testuitslagen voor diverse subpopulaties (met name zieken en niet-zieken), dan alleen over het onderscheidend vermogen bij een bepaald afkappunt (hoofdstuk 20). Het waarden- c.q. utiliteiten-aspect kan men dan desgewenst zelf toevoegen. De afwegingen ten aanzien van het nut van het verrichten van diagnostisch onderzoek of van het volgen van bepaalde strategieën kunnen anders uitvallen voor huisarts en specialist (hoofdstuk 13). Dit houdt in dat diagnostische protocollen 'echelon-specifiek' moeten worden ontwikkeld en gedoceerd. In het kader van die protocolontwikkeling is ook een kritische houding ten opzichte van het therapeutische aanbod van de tweede lijnsgeneeskunde nodig. Zo zal een protocol-aanbeveling voor het handelen naar aanleiding van het vinden van een carotissouffle sterk vereenvoudigd worden na bestudering van de literatuur over de effectiviteit van carotis operaties (hoofdstuk 18). De vraag doet zich zelfs voor: 'is it safe to auscultate the neck?' (13).

Het objectief bepalen en subjectief schatten van a priori kansen op ziekte zijn belangrijke pijlers van de medische besluitvorming. Gezien

de verschillen tussen eerste en tweede lijn, die ook doorklinken in subjectieve schattingen (hoofdstuk 15), is dit een belangrijk studie-terrein voor huisartsen. Vanwege beperkingen van klachten- en morbiditeits-registraties zal het daarbij vaak nodig zijn om meer gericht en meer gestandaardiseerd onderzoek te verrichten (hoofdstuk 14). Ook onderlinge uitwisseling en bespreking van gehanteerde a priori kansen zijn zinvol als vorm van intercollegiale vergelijking (hoofdstuk 16). Men moet ervoor waken dat de informatie die besloten ligt in de a priori kans en de informatie welke de testuitslag in zich draagt, niet door elkaar gaan lopen. In dat geval is de test geen onafhankelijke toetssteen meer. Er is in dit verband veel voor te zeggen dat artsen in eerste instantie dié informatie over de patiënt aan een geraadpleegde diagnost verstrekken die nodig is voor een technisch verantwoorde uitvoering van de test, zonder daarbij al te sturen in de richting van de eigen vermoedens. Men kan eventueel afspreken dat wel een differentiaaldiagnose verstrekt wordt, doch in willekeurige volgorde en zondere nadere aanduidingen van waarschijnlijkheid. Het gaat er om, dat testuitslag en klinische interpretatie verschillende zaken zijn. (hoofdstuk 19).

Besliskundige analyses kunnen veel inzicht geven in de mogelijke consequenties van de diverse opties in termen van goede en minder goede uitkomsten (inclusief neveneffecten en complicaties), op grond van de bestaande stand van kennis. Dit kan geschieden (14):

- kwantitatief: als er voldoende getalsmatige data beschikbaar zijn, en als het mogelijk is op redelijke wijze utiliteiten te bepalen c.q. hun invloed na te gaan;
- kwalitatief: men beperkt zich dan tot het structureren van het probleem, en kan hieruit vaak al conclusies trekken over de zin en onzin van bepaalde (tussen)stappen.

Deze benaderingen kunnen hun vruchten afwerpen ten aanzien van het onderwijs, de nascholing, de intercollegiale toetsing, en ter ondersteuning van consensusbesprekingen (14). Voor de individuele patiëntenzorg zullen de toepassingsmogelijkheden voorsnog beperkt zijn, omdat te weinig kwantitatieve gegevens bekend zijn voor bijzondere situaties. De problemen die men ontmoet bij het verrichten van besliskundige analyses, liggen niet zozeer op het gebied van de rekentechnieken, maar betreffen veel meer (15-19):

- de structurering van het probleem: het is moeilijk om alle relevante strategieën met een bevredigende mate van detaillering in een beslissboom weer te geven;
- de kansen: veel data ontbreken, of kunnen slechts worden ontleend aan weinig overtuigende studies. Men moet bedacht zijn op diverse vormen van 'publicatiebias'. Bovendien is een vraag op welke wijze men data die afkomstig zijn uit verschillende studies op één noemer brengt, respectievelijk kan en mag combineren. Vaak is ook niet bekend in hoeverre diverse kansen van elkaar afhankelijk zijn. Men kan ontbrekende data aanvullen door subjectieve schattingen van deskundigen, maar men moet zich dan goed bewust zijn van de mogelijke bronnen van vertekening;
- utiliteiten: voor de vaststelling van de 'waarde' van diverse gezondheidstoestanden bestaan verschillende technieken, waarbij ook de beleving en de voorkeuren van de patiënt geïncorporeerd worden (20). Niettemin zijn de resultaten hiervan, zeker als het om kwaliteit van leven gaat, steeds vatbaar voor discussie; en bijvoorbeeld afhankelijk van het tijdstip waarop, de toestand waarin, en de techniek waarmee de patiënt ondervraagd wordt (16-18).

Vooraf met betrekking tot de huisartsgeneeskunde zijn deze moeilijkheden niet gering: men werkt immers vaak met weinig voorgestructureerde problemen, en 'zachte' uitkomsten. Over kansen is zeer weinig bekend. Al deze kanttekeningen kan men echter ook maken ten aanzien van de dagelijkse gezondheidszorg waar men, zij het impliciet, voortdurend problemen structureert, kansen percipieert en uitkomsten waardeert. Hierbij vindt veel onder- en overschatting van kansen plaats, en de voorkeuren van de patiënt en de waardering van kwaliteitsaspecten spelen vaak een bescheiden rol. Besliskunde onthult de bestaande controversen en onzekerheden en stelt ze expliciet aan de orde, en vervult daarmee wellicht haar belangrijkste functie (21). De besliskunde is dan ook bij uitstek geschikt om hiaten op te sporen, en daarmee aangrijpingspunten te leveren voor nader onderzoek. Tevens noodzaakt zij telkens tot het expliciteren van vooronderstellingen, in plaats van het routinematig toepassen van eenmaal geleerde 'clinical rules' (22).

## 5. Aanbevelingen

Ten aanzien van het klinisch-epidemiologisch onderzoek op het gebied van klachten, symptomen en diagnostica in de eerste lijn, benadruk ik de volgende aandachtsvelden:

### 1. Theorievorming.

Theorievorming is met name nodig ten aanzien van de methodologie van het huisartsgeneeskundig onderzoek. Enerzijds zullen bruikbare principes en technieken van wetenschapsgebieden met een sterkere onderzoekstraditie (klinische en biomedische wetenschappen, epidemiologie, gedragswetenschappen) moeten worden gemobiliseerd. Anderzijds dienen de specifieke methodologische vraagstukken van het huisartsgeneeskundig onderzoek in kaart te worden gebracht. Gedacht kan hierbij worden aan de bestudering van:

- het probleem van de 'gouden standaard' van de morbiditeit in de eerste lijn;
- uitkomstmaten voor gezondheid, welzijn, ziekte en onwelzijn (23,24);
- mogelijkheden en beperkingen van registratiesystemen in de eerste lijn;
- selectiebias en classificatiebias bij morbiditeitsonderzoek;
- selectiebias ten gevolge van verwijzing. Hierbij komt ook de verhouding tussen de eerste en tweede lijn aan de orde, wat betreft de wederzijdse generaliseerbaarheid van de kennis in beide sectoren;
- methodologische, wetenschapstheoretische en ethische aspecten van de toepassing van besliskunde in de huisartspraktijk.

### 2. Empirisch onderzoek.

- Onderzoek naar het verband tussen klachten/symptomen en aandoeningen in de eerste lijn, en het op grond van empirisch onderzoek documenteren van de selectiebias ten gevolge van verwijzing;
- Het bestuderen van de waarde van de anamnese in de huisartspraktijk voor de diagnostiek; de voorspellende waarde van klachten en symptomen;

- Onderzoek naar de validiteit van (oude en nieuwe) diagnostica in de huisartspraktijk. Hierbij dient er speciale aandacht te zijn voor diagnostica met een "breed screenend" karakter (inspelend op de grote probleemruimte bij de huisarts (25)), diagnostica die eenvoudig toepasbaar (te maken) zijn, en diagnostica die een goed onderscheidend vermogen hebben bij nog weinig ontwikkelde ziektestadia;
- Identificatie van die diagnostische gegevens die voor de huisarts van belang zijn voor het formuleren van een effectieve therapeutische strategie;
- Het evalueren van cruciale, doch nog onvoldoende getoetste onderdelen van diagnostische protocollen door middel van besliskundige analyse en gericht onderzoek;
- Risicoschatting en prognostiek op grond van epidemiologische onderzoeksresultaten, en de mate waarin, bij gebrek aan objectieve gegevens, subjectieve schattingen en consensus hierop een aanvulling kunnen zijn.

#### LITERATUUR.

1. Gerritsma JGM, Smal JA. De werkwijze van huisarts en internist. Bunge, Utrecht 1982.
2. Es JC van, Gerritsma JGM, Smal JA. Simulatie van patiënten in het onderwijs. Instituut voor Huisartsgeneeskunde, Utrecht, 1977.
3. Eijk JThM van. Verschillen in praktijkvoering van huisartsen. Medisch Contact 1979; 34: 987-991.
4. Lamberts H. Hoe beïnvloedbaar is de huisarts? Huisarts en Wetenschap 1981; 24: 371-378 en 424-430.
5. Barr DM, Norem DG. Factors influencing clinical decision making by physicians: a conceptual framework and literature review. Office of Community Health Research, Rockford School of Medicine, University of Illinois, College of Medicine, 1978.
6. Holten-Vriesema J, Tompot C, van Aalderen H et al. Methodisch werken, over een algemene methode van hulpverlening en de opbouw van een functionele relatie, toegespitst op de huisartspraktijk. Huisarts en Wetenschap 1978; 21: 322-335.

7. Knottnerus JA. Democratisering in praktijk, dilemma's voor de hulpverleners. Tijdschrift voor gezondheid en politiek 1983; 1, december: 6-11.
8. Knottnerus JA, de Vet HCW. Effectiviteitsonderzoek in de huisarts-praktijk. Aanvaard voor publicatie in Huisarts en Wetenschap.
9. Ellenberg JH, Nelson KB. Sample selection and the natural history of diseases: studies of febrile seizures. JAMA 1980; 243: 1337-1340.
10. Weel C van. Alledaagse ziekten en een gezond verstand. Inaugurele rede, Nijmegen 1986. NUHI, Nijmegen 1986.
11. Lamberts H. Redenen om naar de huisarts te gaan. Eerste ervaringen met de Reason for Encounter Classification. Huisarts en Wetenschap 1982; 25: 301-310.
12. Velden HGM van der. Diagnose of prognose; de betekenis van de epidemiologie voor het handelen van de huisarts. Huisarts en Wetenschap 1983; 26: 125-128.
13. Kuller LH, Sutton KC. Carotid artery bruit: is it safe and effective to auscultate the neck? Stroke 1984; 15: 944-947.
14. Everdingen JJE van, Touw PPJ, Klazings NS, Casparie AF. Het belang van consensus in de geneeskunde. Ned Tijdschr Geneesk 1985; 129: 1911-1913.
15. Schwartz WB, Gorry GA, Kassirer JP, Essig A. Decision Analysis and clinical judgment. Am J Med 1973; 55: 459-472.
16. Zarin DA, Pauker SG. Decision analysis as a basis for medical decision making: the tree of Hippocrates. J Medicine and Philosophy 1984; 9: 181-213.
17. Elstein AS, Holmer MM, Ravitch MM, Rovner DR, Holzman GB, Rothert ML. Medical Decisions in perspective: applied research in cognitive psychology. Perspectives in Biology and Medicine 1983; 26: 486-499.
18. McNeil BJ, Pauker SG. Decision Analysis for public health: principles and illustrations. Ann Rev Public Health 1984; 5: 135-161.
19. Feinstein. The haze of Bayes, the aerial palaces of decision analysis, and the computerized Oija board. Clinical Pharmacology and Therapeutics 1977; 21: 482-496.
20. Weinstein MC, Fineberg HV. Clinical Decision Analysis. WB Saunders Philadelphia 1980.



21. Sturmans F, Verbeek ALM, van Daal WAJ. Protocolaire geneeskunde: dictatuur of verantwoord kiezen? In: Het Medisch Jaar 1985. Bohn, Scheltema en Holkema, Utrecht/Antwerpen 1985; p 395-407.
22. Eddy DM. Clinical Policies and the quality of clinical practice. *New Eng J Med* 1982; 307: 343-347.
23. Wulff HR. Current trends in clinical thinking. *Acta Med Scand* 1983; 213: 3-4.
24. Donabedian A, Wheeler JRC, Wyszewianski L. Quality, Cost and Health: an integrative model. *Medical Care* 1982; 20: 975-992.
25. Weel C van. Achilles tot hardloper bekeerd. *Huisarts en Wetenschap* 1981; 24: 417-424.

---

SAMENVATTING.

In dit proefschrift wordt op verschillende wijzen en aan de hand van diverse deelonderwerpen de interpretatie van diagnostische informatie bestudeerd, vanuit een epidemiologische invalshoek. Speciale aandacht wordt besteed aan de relatie tussen eerste en tweede lijn.

In deel I worden algemene begrippen besproken, die van nut zijn bij het analyseren van de waarde van diagnostische informatie. Daarbij komt aan de orde, welke gevolgen selectie door middel van verwijzing kan hebben. Hoofdstuk 2 benadrukt dat verschillen in de prevalentie van aandoeningen tussen de huisartsgeneeskundige en de specialistische praktijk consequenties hebben voor de interpretatie van diagnostische gegevens en de te volgen diagnostische strategie. In hoofdstuk 3 wordt gedemonstreerd dat in patientenpopulaties die geselecteerd zijn door middel van verwijzing een ander beeld kan ontstaan van verbanden tussen klachten en bevindingen dan in ongeselecteerde populaties. Als men bij het bestuderen van deze verbanden geen adequate vergelijkingsgroepen kiest kunnen schijnverbanden ten onrechte steeds weer bevestigd worden.

Deel II bevat de verslaglegging van een aantal onderzoeken naar het verband tussen veelvoorkomende klachten en diagnostische bevindingen. Op grond van gegevens uit een bevolkingsonderzoek onder Amsterdammers, waarbij de bloeddruk werd gemeten en gevraagd werd naar het gebruik van middelen tegen hoofdpijn, werd het verband tussen hoofdpijn en hoge bloeddruk bestudeerd (hoofdstuk 4). Bij personen die wekelijks hoofdpijnmiddelen gebruiken komt een diastolische bloeddruk van 105 mmHg of hoger iets vaker voor (5 % versus 3 %). Tegelijkertijd blijkt echter, dat het last hebben van hoofdpijn nauwelijks discrimineert tussen hoge en lagere bloeddrukken. Er zijn, kortom, voor patient en arts weinig argumenten om het meten van de bloeddruk af te laten hangen van hoofdpijnlachten.

Vervolgens komt een andere "vage" klacht aan de orde: in hoeverre hebben moeheidsklachten, gepresenteerd aan de huisarts, voorspellende betekenis ten aanzien van het hebben van bloedarmoede of ijzer gebrek? Deze vraag is onderzocht door middel van "diagnostisch patient-contro-

le-onderzoek". Bij patienten die geselecteerd waren via verwijzing naar een diagnostisch centrum werd over het geheel genomen geen verband gevonden tussen moeheid en Hb-gehalte. Alleen bij oudere mannen was er een trend in de verwachte richting (hoofdstuk 5). In een niet door verwijzing geselecteerde populatie kon ook de laatstgenoemde trend niet worden teruggevonden. Dit onderzoek richtte zich, ter vermindering van "ruis", op patienten met onverklaarde moeheidsklachten - al dan niet in combinatie met andere "anemische" klachten - die door de huisarts als indicatie werden beschouwd voor hematologisch onderzoek (hoofdstuk 6). Ook tussen deze klachten en de ijzervoorraad, gemeten door middel van het serumferritine, werd geen relatie gevonden (hoofdstuk 7). Zeer lage Hb-waarden kwamen bij deze patienten niet voor. Toch werd, onafhankelijk van klachten en Hb-gehalte, bij 11 % van de patienten een serumferritine van 12 g/l of lager gevonden. Omdat zwangeren van de zojuist besproken studies waren uitgesloten, en omdat hematologisch onderzoek juist bij hen veel wordt verricht, werd het verband tussen (moeheids- en andere) klachten en hematologische parameters apart bestudeerd bij zwangeren. Gekozen werd voor een screeningsonderzoek rond de dertigste zwangerschapsweek met een vragenlijst en hematologische bepalingen in drie subpopulaties: vrouwen die door perifere verloskundigen werden begeleid, en vrouwen die op een polikliniek verloskunde werden begeleid door verloskundigen respectievelijk gynaecologen. Nu werden wél samenhangen gevonden: met name de klachten "moeheid, erger dan vóór de zwangerschap", kuitkrampen, hartkloppingen en duizeligheid wijzen op een lager Hb en/of Ht. Omdat het verband tussen klachten en een laag Hb vooral duidelijk is bij hogere MCV-waarden, zou foliumzuurtekort hierbij een rol kunnen spelen (hoofdstuk 8).

Via beschrijvend follow-up onderzoek is nagegaan welke diagnostische eindconclusies de huisarts trok naar aanleiding van een tweetal klachten: onverklaarde moeheid, en klachten op de borst (hoofdstukken 9 en 10). Moeheidsklachten leiden, ook in tweede instantie, meestal niet tot een diagnostische conclusie, en dit geldt vooral voor jongere personen. Bij klachten op de borst is twijfel ten aanzien van een eventuele cardiale oorsprong, ook nog na een half jaar, een factor van betekenis. Beide studies schetsen een beeld van de onzekerheden waarmee huisartsen werken, en de mate waarin via verwijzing geselecteerd

wordt.

Deel III bespreekt en illustreert de medisch besliskundige benadering, wederom met speciale aandacht voor de verhouding tussen eerste en tweede lijnsgeneeskunde. Nadat een oriënterend overzicht gegeven is (hoofdstuk 11), wordt benadrukt, dat het voor de beoordeling en (her)-waardering van de zeefunctie van de huisarts essentieel is dat men overzicht heeft over "alle vier velden" van de huisartsgeneeskundige betrokkenheid ten aanzien van verwijzing (hoofdstuk 12). Aan de hand van een voorbeeld (de palpabele lymfklier) wordt verduidelijkt dat men er bij het formuleren van diagnostische protocollen mee rekening moet houden dat de aanbevelingen voor huisartsen en specialisten er geheel anders uit kunnen zien. De a priori kans op ziekte, al dan niet beïnvloed door selectie, speelt hierbij een sleutelrol (hoofdstuk 13). Dit brengt ons op de vraag naar de betekenis van het begrip a priori kans in de huisartsgeneeskunde en op de wijze waarop hiervan een empirisch onderbouwd beeld kan worden verkregen. Routinematig verzamelde registratiegegevens hebben wat dit betreft beperkingen, afhankelijk van het type aandoening waarin men geïnteresseerd is (hoofdstuk 14). Dat subjectieve schattingen van a priori kansen van huisartsen en specialisten kunnen verschillen, blijkt uit hoofdstuk 15: naar aanleiding van een papieren patiënte met moeheidsklachten nemen internisten vooral anemie en maligniteit in hun differentiaal-diagnose op; Bij huisartsen staat depressie voorop, gevolgd door anemie. Ook zijn er verschillen wat betreft het beleid: internisten doen meer onderzoek dan huisartsen. Ook tussen huisartsen (-in opleiding) kunnen grote verschillen bestaan: voor een patiënt met klachten van vermagering liepen de schattingen van de kans op kanker uiteen van 10 % tot 55 %. Het "niet-pluis gevoel" hing meer samen met de angst voor somatische aandoeningen dan met de angst voor meer functionele stoornissen (hoofdstuk 16). Diagnostiek wordt gelegitimeerd door haar belang voor therapeutische beslissingen en de gunstige beïnvloeding van de prognose. Als een therapie niet effectief is, is diagnostiek gericht op de indicatiestelling voor die therapie niet zinvol (hoofdstuk 17). In dit verband is de vraag gesteld: moet er naar aanleiding van de bevinding "asymptomatische carotissouffle" nader onderzoek plaatsvinden ter vaststelling van (de graad van ernst van) een locale stenose en de operabili-

teit ervan? Bij gebrek aan gerandomiseerd onderzoek moet men zich voor de beantwoording van deze vraag voorlopig baseren op gepubliceerde patientenseries. Om deze systematisch te kunnen verwerken, verrichtten wij een besliskundige analyse. De conclusie is dat de literatuur geen ondersteuning biedt voor een op operatie gericht beleid (hoofdstuk 18).

Ter afsluiting van deel III wordt de aandacht gevestigd op de wenselijkheid van het onderscheiden van "testuitslagen sec" en de interpretatie ervan in het licht van de a priori informatie (hoofdstuk 19).

Deel IV gaat nog wat dieper in op de betekenis van diagnostische tests. Diagnostiek kan met diverse oogmerken worden verricht. Hiermee hangen samen de verschillende betekenissen van de begrippen "normaal" en "afwijkend". Hiervan wordt in hoofdstuk 20 een inventarisatie gegeven. De vraag of een bepaalde test uitslag nu "afwijkend" is of niet is niet in het algemeen te beantwoorden. Uiteindelijk is er sprake van een besliskundige probleemstelling, waarbij de waarden van de mogelijke uitkomsten worden beoordeeld in relatie tot het gestelde doel. Daarbij staat ook het nut van het testen zelf ter discussie. Bij het verstrekken van referentiewaarden is het aan te bevelen de volledige distributies van uitslagen in de referentiepopulaties te geven, en niet slechts de waarden van één of enkele afkappunten.

Het onderscheidend vermogen van test is gevoelig voor selectie door middel van verwijzing (hoofdstuk 21). Niet alleen de prevalentie, maar ook de test-kenmerken worden beïnvloed. Temidden van de toename van de belangstelling voor het - vaak als vast gegeven opgevatte - onderscheidend vermogen van diagnostische tests, dient men zich van deze complicaties bewust te zijn en de reikwijdte ervan te onderzoeken.

In deel V worden twee onderzoeken beschreven naar de prognostische betekenis van uitslagen van diagnostische i.c. hematologische tests. Middels een cohortstudie werd het verband bestudeerd tussen het Hb en de hematocriet van zwangeren enerzijds en het later optreden van prematuritas en een laag geboortegewicht anderzijds (hoofdstuk 22). Ook na correctie voor andere belangrijke covariaten (roken, pariteit, bloeddruk, gebruik van ijzermedicatie), blijkt met name een hoge hematocriet positief samen te hangen met een minder gunstige afloop.

Mede op grond van de eerder beschreven samenhang van een laag Hb/Ht met enkele subjectieve klachten, bevelen wij aan tijdens de zwangerschap te streven naar een hematocriet die, na de fysiologische hemodilutie, tussen de 32% en 38 % bedraagt.

In hoofdstuk 23 wordt via een patient-controle onderzoek bij mannelijke werknemers nagegaan of er een verband is tussen matig verhoogde Hb- en Ht-waarden en de kans op het krijgen van een hartinfarct. Dit blijkt inderdaad het geval, ook rekening houdend met andere relevante factoren.

Deel VI bevat een afsluitende beschouwing waarin deelonderwerpen in onderling verband worden besproken. In dit verband wordt een indeling voorgesteld van het werkveld van de klinische epidemiologie. Tenslotte worden een aantal aanbevelingen gedaan ten aanzien van het klinisch-epidemiologisch onderzoek op het gebied van de diagnostiek in de huisartspraktijk.

## SUMMARY.

This thesis deals with the interpretation of diagnostic information, from a clinical epidemiological point of view. Theoretical backgrounds are discussed and practical examples are presented, with a special interest in the relationship between primary and specialist care.

In part I basic concepts, of importance when analyzing diagnostic information, are introduced. The consequences of selection by referral are stipulated. Chapter 2 explains the impact of different prevalences with respect to the interpretation of diagnostic data and the planning of diagnostic strategies. Chapter 3 demonstrates the changes in the relationship between symptoms and diagnostic findings, that may occur as a result of referral. Pseudo-relationships can be maintained by erroneous comparisons.

Part II contains the results of a number of studies of the relationships between common complaints and diagnostic findings.

Using data from a population survey in Amsterdam, the relationship between headache and high bloodpressure was investigated (chapter 4). Among subjects who took anticephalalgics weekly, a diastolic pressure of 105 mmHg or higher is a little more frequent (5% versus 3%).

However, suffering headache hardly discriminates between high and low bloodpressures. Bloodpressure surveillance, guided by headache, is not supported by these results. Fatigue, as an indicator of anemia, is investigated in two "diagnostic case-control studies". Among patients referred to a diagnostic centre no overall relationship between fatigue and hemoglobin level (Hb) was demonstrated. Only in older men a trend in the expected direction was found (chapter 5). Such a trend was not confirmed in a general practice study among patients not preselected by referral. For efficiency reasons, this study focused on "unexplained" fatigue considered to be an indication for hematological evaluation (chapter 6). Also between this complaint and the iron stores, as estimated by the serumferritin level, no relationships were observed. Extremely low Hb-values have not been found among these patients. Yet, 11% of them had serumferritin levels of 12 mg/l or lower (chapter 7).

Pregnant women had been excluded from the studies just mentioned, whereas their hematological status is examined routinely. Therefore, the relationship between symptoms and hematologic parameters among pregnant women was studied separately. The women were screened at about 30 weeks of gestation with blood examinations and a questionnaire. This time, moderate relationships were found: increased fatigue, calf cramps, palpitations and dizziness are indicators of low Hb and/or Ht-values. Because the relationship between symptoms and Hb-level is most clear in the presence of a higher Mean Cellular Volume (MCV), a possible role of folic acid deficiency is suggested (chapter 8).

In two descriptive studies, the general practitioners' diagnostic conclusions in connection with unexplained fatigue and chestpain are established (chapters 9 and 10). In most cases of fatigue no diagnostic conclusion is reached, and this is the more so in younger patients. Doubt about a cardiac origin of chestpain is a factor of importance, even half a year after first consultation. Both studies give an impression of the uncertainties general practitioners are dealing with in daily practice, and of the selection by referral.

Part III discusses and illustrates clinical decision analysis, again with a special interest into the relationship between primary care and specialist care. After a general introduction (chapter 11), it is stressed that in order to (r)evaluate the selective function of the general practitioner, it is necessary to have a representative overview over the "four fold-table" of his clinical experience (chapter 12). By means of an example (palpable lymph node) it is made clear that optimal protocols may not be the same for general practitioners and specialists. Prior probabilities have a great impact (chapter 13). The significance of prior probabilities in general practice and the methods to estimate them are discussed. Registration systems have limitations, depending upon the type of symptom and disease of interest (chapter 14). Subjective estimations of prior probabilities by general practitioners and specialists may differ. Considering exactly the same case, i.e. a fatigue patient, in differential diagnosis internists give much weight to anemia and malignancy. General practitioners however think depression is the most likely diagnosis,



followed by anemia. Furthermore, internists perform more diagnostic investigations (chapter 15). Also, between general practitioners (residents) considerable differences are observed: for a patient with weight loss the estimated probabilities of malignancy varied from 10 to 55%. Feeling, that something is 'wrong', appears to be more related with fear of somatic pathology, than with fear of functional problems (chapter 16).

Diagnostics are legitimated by their impact on management decisions in order to favour prognosis. Diagnostic investigations to establish the indication for an ineffective therapy are useless and redundant (chapter 17). In this connection we considered the question: is a asymptomatic carotid bruit a reason for further investigations to evaluate operability? Because randomized trials are lacking, the answer must be derived from published patient series. Using their results in a systematic fashion, we performed a clinical decision analysis. We concluded, that an operative strategy is not supported by literature data (chapter 18).

Finally, attention is drawn to the need to distinguish between 'test-results sec' and their interpretation in the light of prior information.

In part IV, the interpretation of diagnostic data is discussed more deeply. Different meanings of 'normal' and 'abnormal' are associated with the aims of testing. Chapter 20 provides an overview. Eventually, the search for reference values leads to a decision analytic approach, taking into consideration the expected utilities of each possible outcome. Also, the performance of the test itself is included in the decision. In presenting reference values it is recommended to provide the whole distribution of test results in the compared populations, and not only the arbitrarily chosen cut-off points. Discrimination of a test is sensitive to selection by referral patterns (chapter 21). Both prevalences and test characteristics are influenced. This notion is important, because the interest in test characteristics is growing presently, and the latter often are supposed to be constants.

Part V describes two studies concerning the prognostic significance of test results, i.c. hematological parameters. In a cohortstudy the

relationships between Hb and Ht in pregnancy, and preterm birth and low birth weight respectively, were investigated (chapter 22). After adjustment for smoking, parity, bloodpressure and iron medication, a high hematocrit especially appears to be associated with an adverse outcome. In connection with the earlier described relationship between low Ht values and subjective symptoms, we advise a reference range of the Ht from 32 to 38% during the hemodilution phase in pregnancy. Chapter 23 describes a case-controlstudy in male workers to examine the relationship between moderately high Hb- and Ht-values and the occurrence of myocardial infarction. Positive relationships are found, also after adjustment for covariates.

In part VI the contents of this thesis are discussed in an integrant framework. A simple taxonomy of clinical epidemiologic studies is proposed. Recommendations are made with respect to diagnostic epidemiological research in general practice.



### Dankwoord.

Aan de totstandkoming van de diverse onderdelen van dit proefschrift hebben enkele duizenden personen een bijdrage geleverd. Daarmee is in de praktijk gebracht, dat epidemiologie met grote groepen werkt. Dat gold in dit geval niet alleen voor hen die onderzocht werden ( $n=5868$ ), maar ook voor degenen die het onderzoek verrichtten ( $n \geq 58$ ). Ten aanzien van hen allen spreek ik mijn erkentelijkheid uit. Dit dankwoord betreft behalve de medewerking zelf vooral ook de plezierige wijze waarop deze gestalte kreeg. Het is niet als afsluiting bedoeld: ik hoop dat er vanuit de opgebouwde contacten en ervaringen nog veel meer tot stand zal komen.

Een aantal personen wil ik op deze plaats in het bijzonder bedanken. Om te beginnen, ben ik Prof.dr. Paul Knipschild en Prof.dr. Ferd Sturmans dankbaar voor het vertrouwen dat zij in mij hebben gesteld. Paul Knipschild heeft mij geleerd dat het voor het doen van onderzoek nodig is om, soms pijnlijke, reducties aan te brengen, omdat ook hier geldt dat zachte heelmeesters stinkende wonden maken. Zijn wetenschappelijke creativiteit stond garant voor een aantal jaren van boeiende samenwerking. Paul, de wijze waarop je met mij in de theorie en de praktijk rond dit proefschrift bent gedoken, is voor mij van grote betekenis geweest.

Ferd Sturmans heeft met zijn inspirerende publicaties mijn belangstelling voor het vak epidemiologie gewekt, reeds lang voordat onze paden elkaar kruisten. Hij heeft mij, en niet alleen mij, gebracht tot het belangrijke inzicht, dat vanuit de epidemiologie algemene begrippen als medicalisering en iatrogenese geoperationaliseerd en gewogen kunnen worden in het kader van specifieke medische probleemstellingen. Ferd, bedankt voor de onmisbare steun en feed-back die ik van je kreeg.

In één adem met de zojuist genoemden wil ik Ir. Pieter Leffers dankzeggen. Pieter, alleen al door vier jaar je kamergenoot te zijn geweest heb ik zoveel van je opgestoken dat ik verwacht daaruit nog gedurende een veelvoud van vier jaren te kunnen putten.

Een speciaal dankwoord wil ik uitspreken ten aanzien van hen die aan onderdelen van dit proefschrift een bijzondere bijdrage hebben geleverd.

Dr. Jan van Wersch, Arthur Sijstermanns, Ad Visser, Richard Starmans en Jos Didden hebben sleutelrollen vervuld in het onderzoek in Heerlense huisartspraktijken. Edzo Ebbens en Phile Govaert deden dat in het onderzoek, verricht in het kader van het Registratienet Huisartspraktijken. Van cruciaal belang in deze beide onderzoeken was de actieve deelname van de huisartsen, waarvoor ik hen zeer erkentelijk ben.

Zonder de inzet van Luis Delgado, Truus Stenkens en Myra Pereboom en de steun van Frans Smits en Gerard Essed was het onderzoek bij zwangeren niet mogelijk geweest. Speciale vermelding verdienen in dit verband de eerstelijns-verloskundigen Mevrouw J.W.M. Heuts-verstraten, Zr. Bernardi Kok, Mevrouw A.C.M. Renkens-Gommans, en Mevrouw M. Lamkin. Zij hebben gedurende enkele maanden zelfs hun wijze van bloedafname afgestemd op de behoeften van het onderzoek.

Met Gerard Swaen, Jos Slangen, John Durinck en Lex Volovics heb ik met plezier samengewerkt aan het onderzoek naar de relatie tussen hemoglobinegehalte en myocardiinfarct.

Annette Boenink en Paul de Vos wil ik bedanken voor hun in dit proefschrift geëtaleerde werkzaamheden in het kader van hun student-assistentenschap.

Piet Kerkhof heeft de data van het Diagnostisch Centrum Maastricht ter beschikking gesteld en bewerkt. Dr. Jan Broekman, Dr. Jan Lodder en Dr. Herman Muller hebben wezenlijke bijdragen geleverd aan hoofdstukken in dit proefschrift.

Riny Bodifée, Phia Breevoort en Thum Aarts hebben de totstandkoming van het proefschrift letterlijk begeleid. Dit moet een niet geringe opgave zijn geweest, en om daarvoor te danken schieten woorden te kort.

Hartelijke dank ben ik ook verschuldigd aan Trudie Seegers en Jim Tatipata, die mij in de laatste fase geholpen hebben met de tekstverwerking en zich daarbij niet stoorden aan vreemde tijdstippen en pechduiveltjes.

Tenslotte ben ik Prof.dr. Cees de Geus en Prof.dr. Henk van der Velden zeer erkentelijk voor hun opbouwende commentaar bij de voorbereiding van dit proefschrift.

Curriculum vitae.

Johannes Andreas Knottnerus werd op 4 september 1951 geboren te Nieuw-Beerta. In 1969 behaalde hij het eindexamen gymnasium bèta aan het Baarnsch Lyceum, waarna hij geneeskunde studeerde aan de Vrije Universiteit te Amsterdam. Na het artsexamen in 1977, volgde hij aan dezelfde instelling de huisartsopleiding. Vervolgens was hij in Amsterdam werkzaam als huisarts en als wetenschappelijk medewerker aan het Huisartsen Instituut van de Vrije Universiteit. Sinds 1982 is hij verbonden aan de Rijksuniversiteit Limburg, aanvankelijk als wetenschappelijk medewerker bij de capaciteitsgroep Epidemiologie, en vanaf medio 1986 als universitair docent bij de vakgroep Huisartsgeneeskunde.